AFIT/GE/ENG/96D-12

OPTIMIZATION OF A GPS-BASED NAVIGATION REFERENCE SYSTEM

THESIS

Jason Bryce McKay
Second Lieutenant, USAF

AFIT/GE/ENG/96D-12

PII Redacted

AFIT/GE/ENG/96D-12

OPTIMIZATION OF A GPS-BASED NAVIGATION REFERENCE SYSTEM

THESIS

Presented to the Faculty of the School of Engineering
of the Air Force Institute of Technology
Air University
In Partial Fulfillment of the
Requirements for the Degree of
Master of Science in Electrical Engineering

Jason Bryce McKay, B.S.E.E.
Second Lieutenant, USAF

December, 1996

Approved for public release; distribution unlimited

Acknowledgements

First and foremost I would like to thank my wife, Kristen, for marrying me anyway, knowing she would have to share me with this thesis. I don't know what I would have done without her love and support during these long months at AFIT.

My thesis advisor, Dr. Meir Pachter, deserves special thanks for his support and enthusiasm throughout the course of this thesis. His experience and insight helped me to stay focused on my research, and his perpetual good spirits and enthusiasm helped me to enjoy it.

Thanks must also go out to Lt. Nathan White and Capt. Bob Lewis, my fellow guidance and control classmates, for all the things they did for me. It would have been pretty dull in the navigation lab without them. I would also like to thank Mr. Donald Smith, the resident computer wizard, for putting up with my questions, and rescuing the network from countless user mistakes.

Finally, I would like to thank my family. I would not be where I am today without their support and advice. Thanks for everything.

Jason Bryce McKay

Table of Contents

List of Figures

Figure		Page

Figure Page

AFIT/GE/ENG/96D-12

Abstract

The integration of Global Positioning System (GPS) measurements into the navigation systems of Air Force aircraft has enhanced their accuracy. These increasingly accurate new aircraft navigation systems require the Air Force to continually improve its navigation reference system in order to accurately test them. The new Air Force navigation reference system under development, the Submeter Accuracy Reference System (SARS), will achieve the necessary accuracy with carrier phase GPS technology. The SARS is an inverted GPS system which consists of an array of GPS receivers on the ground and an airborne pseudolite mounted on the aircraft whose navigation system is to be tested. The SARS will provide a proof position estimate that is used to check the navigation system under test. Unfortunately, ground based inverted GPS systems such as the SARS tend to suffer from high geometric sensitivity to measurement errors. The SARS has the potential to attain the level of accuracy required, provided the measurement geometry is chosen to minimize the SARS' sensitivity to measurement errors. This research tackles the problem of optimizing the SARS' receiver array configuration to minimize the system's sensitivity to pseudorange errors, thus enhancing the position reference produced. The analysis determines that the proper choice of cost function for the optimization is the condition number of the visibility matrix H, rather than the commonly used GDOP. Insight into the problem is provided by a graphical technique for evaluating receiver array geometry. Moreover, two receiver array numerical optimization programs are developed. The results of the receiver array optimization programs show that the geometric sensitivity to error in the SARS airspace can be reduced to acceptable levels through proper array design. Several good receiver array designs are shown. Finally, a technique for further reducing the geometric sensitivity of the SARS is discussed.

OPTIMIZATION OF A GPS-BASED NAVIGATION REFERENCE SYSTEM

I. *Introduction*

1.1 Background

The integration of GPS, a satellite based radio-navigation system, with Inertial Navigation Systems (INS) has produced highly accurate aircraft navigation systems. As the performance of aircraft navigation systems continues to improve, the Air Force must develop a more accurate means of flight testing them. The current method of flight testing an aircraft's navigation system consists of placing a Navigation Reference System (NRS) onboard the aircraft and comparing the performance of the aircraft's navigation system to the performance of the reference system. For this to work, the NRS needs to be significantly more accurate than the system under test, and it needs to be able to fit somewhere onboard the aircraft. The current NRS does not meet either of those two requirements. The current NRS is not significantly more accurate than some of the new systems being tested. It is also too big to be fit onboard many Air Force aircraft, such as fighter and reconnaissance aircraft. Accordingly, the Air Force is developing a new NRS that does satisfy the accuracy and size requirements.

The new Air Force navigation test reference system, under development at the Central Inertial Guidance Test Facility (CIGTF) at Holloman AFB, NM, is the Submeter Accuracy Reference System (SARS). This reference system is especially designed for flight testing integrated navigation systems. The system uses carrier phase GPS technology to obtain the high accuracy needed to test the integrated navigation systems of high performance aircraft. It does not use the GPS satellites, but instead puts a pseudolite (GPS transmitter) on the test aircraft and uses an array of GPS receivers on the ground at surveyed locations. The transmitter and receivers can be set to a different frequency band than the GPS satellites, making the SARS a navigation reference independent of the GPS satellites. Since a pseudolite is small enough to be mounted on any Air Force aircraft, the SARS can be used to flight test any aircraft's navigation system. The SARS has undergone preliminary testing which has validated the concept [17], but did not achieve the the level of accuracy necessary to accurately test the Air Force's newer aircraft navigation systems. It is desired to increase the accuracy of the SARS, to enable it to be used as an accurate proof reference of position.

The desired increase in accuracy of the SARS can be obtained by improving the accuracy of the GPS technology used in the system. As GPS is already highly accurate (and complicated), the factors that impact GPS accuracy and, more importantly, the techniques for further improving GPS accuracy may not be intuitive. Fortunately, there is currently a lot of interest in GPS, so this effort to improve the accuracy of the SARS need not be conducted in a vacuum. There is a large amount of pertinent information in the literature that has been useful in choosing the direction of this research. The following literature review examines the current factors that pertain to improving the accuracy of GPS-based navigators, identifies possible directions of research towards improving the SARS' accuracy, and determines the most promising direction of research.

1.2 Literature Review

The first step in this thesis is to look at the current body of GPS knowledge and design techniques to see what ideas can be found for improving the accuracy of the SARS. This literature review has two main objectives: determining the factors that impact GPS system accuracy and identifying promising directions of research from that information. To put the material into context, the basic concepts of GPS are outlined in the following section. Once the basic concepts are established, the two main objectives are discussed. Within the discussions, the factors affecting GPS accuracy are outlined, their relevance to the SARS' configuration is examined, and their impact on the direction of my research is summarized. Although there is a large body of information on both topics, the ground based configuration of the SARS is sufficiently different from satellite based GPS to cause much of the literature to have limited applicability. Therefore, this review sifts through the current state of GPS knowledge to find what can be applied to my research.

1.2.1 Basic Concepts of GPS. The Global Positioning System (GPS) is a radionavigation system that uses time-difference-of-arrival measurements on signals transmitted from reference stations to determine user position and velocity. In standard, satellite-based GPS, the reference stations are satellites in precisely controlled 12 hour orbits. The satellites transmit time tagged, coded data down to earth, where the user, a GPS receiver in this case, receives the signals and calculates its position. In the SARS, the configuration is reversed. In SARS, the reference stations are receivers at known locations on the ground at the navigation system flight test range, and the user is an aircraft outfitted with a GPS transmitter (a pseudolite). In both cases, the same techniques are used to solve the navigation equations, but the configuration of the SARS is sufficiently different from the satellite-based GPS to make much of the current literature not applicable.

1.2.2 Factors That Affect GPS Accuracy. Many factors affect the accuracy of GPS systems. For discussion purposes, these factors are lumped into four overall groups: the GPS error sources, the system geometry, modeling, and data processing. Although all of these factors affect GPS accuracy, they may have a different impact on the SARS system than on standard GPS, due to the ground-based configuration of the SARS.

1.2.2.1 GPS Error Sources. The many errors that corrupt the GPS navigation solution are a big factor that impacts GPS system accuracy. In an overview of the principles of the GPS system, Milliken et al. list seven basic sources of GPS system errors: satellite (reference station) clock errors, atmospheric delays, group delay, ephemeris (reference station position) errors, multipath errors, receiver noise and resolution, and receiver (user) vehicle dynamics [14]. Since the GPS equipment used to implement the SARS is essentially the same as that used for satellite GPS, it seems likely that the sources of error for the SARS will be similar to those for satellite-based GPS. However, a review of previous AFIT research on the SARS concept indicates that some of the sources of error may be different, or at least of different relative importance for the SARS, because the SARS uses no space based equipment [7].

Significant differences in error sources seem likely in these areas: atmospheric delays, ephemeris errors, multipath errors, receiver noise/resolution, and receiver vehicle dynamics. First, the atmospheric delays will be much smaller for the SARS case, because the distance between the receivers and the transmitters will be very short, compared to the distances involved in satellite GPS. Also, the SARS signals will not have to go through the ionosphere, reducing overall signal distortion. Second, the ephemeris (GPS reference station position) errors will be different, as it is possible to determine the location of ground based receivers (SARS' reference stations) much more accurately than the location of satellites. In fact, the ground based receiver array could use the GPS satellites over a course of days to pinpoint the location of each receiver to the extreme limit of possible accuracy. This quite possibly could all but eliminate ephemeris (reference station position) errors in the SARS. Third, the multipath errors could be different in the SARS. Multipath errors are due to signal reflections, and waveguide effects of the atmosphere that affect signal transit time (different paths with different transit times). In the SARS, reflections off the ground seem likely to dominate the multipath errors, because there are many receivers on the ground (that could receive reflected signals), and the short distances between the transmitter and the receivers would tend to minimize waveguide effects. Fourth, receiver noise and receiver resolution may be much more important to the SARS. Chaffee and Abel, in their work on a closed form solution to the pseudorange (GPS measurement) equations, state that the noise to pseudorange ratio is larger, the closer together the transmitter and receiver are [4]. With the short distances used in the SARS range, this could be a problem. Lastly, the user vehicle dynamics could have a different impact on the SARS than if the satellites were used. This is because the distances between the reference stations and user are so much smaller in the SARS. Changes in the user position that would have no impact on the effective geometry of the configuration if satellites were used could drastically change the effective system geometry of the SARS. As the geometry of the configuration affects the impact the error sources have on the navigation accuracy [7, 14, 17], this problem is critical to the design of the SARS.

Looking at the possible differences in error sources of the SARS, compared to satellite GPS, one can see that the error characteristics of the SARS may prove quite different from satellite GPS, even though the hardware and data processing techniques are common in both systems. This points out that blindly applying standard satellite GPS technology to the SARS, without considering the differences in the systems' error sources, could lead to degraded performance.

1.2.2.2 System Geometry. The second factor that affects GPS accuracy is the system geometry. The geometry of the system means the orientation of the reference stations to each other and to the user. Note that this does depend on the position of the user relative to the array of reference stations, as well as the configuration of the reference station array itself. The geometry of the system has been shown to impact the mapping of measurement errors to position errors in the navigation solution [14,17]. The standard measure in the literature of geometric sensitivity is the Geometric Dilution of Precision (GDOP). This GDOP is a rough measure of the scaling effect that the geometry has on the measurement errors. When the geometry is good, the GDOP and the system's geometric sensitivity to measurement errors are low. When the geometry is poor, the GDOP and the system's geometric sensitivity can become quite high.

In the literature, 'good' geometry (for the satellites) has been related to the three dimensional volume of a tetrahedron formed from the unit vectors drawn from the user position along the directions of the straight line paths between the user and each reference station [14]. This definition of good geometry breaks down when more than the minimum of four reference stations are used at one time, but the gist of the idea is that the more space enclosed by the 'connect the dot' figure formed out of the user and receiver positions, the 'better' the geometry.

In the SARS design, although the geometry has the same impact on GPS accuracy as in the satellite case, there is not nearly as much freedom to choose reference station locations as there is with the satellites. By design, the geometry of satellite GPS is rather good for terrestrial users. This is not so for the SARS. All the reference stations for the SARS will essentially be located on the ground. Looking at the 'good' geometry definition for the satellite case, it would seem that the SARS configuration cannot avoid poor geometry, as was reported in the concept testing [17]. Given that, a major goal of this research is to figure out the best possible configuration for the receiver array, to minimize this problem of 'poor' geometry.

1.2.2.3 Modeling. The next factor that affects GPS system accuracy is the modeling of GPS errors and user dynamics. Through sophisticated filtering and estimation techniques, it is possible to ameliorate the effects of GPS errors and receiver problems due to user dynamics [5]. To do this, statistical models of the GPS errors and vehicle dynamics need to be formed. Much work in this area has been done at AFIT as well as many other institutions, resulting in a proliferation of models for both GPS errors and user dynamics [6,8,16,19,21–23]. Once the models are formed, and are proven in use, they tend to get programmed into the GPS user equipment and processing software. This simplifies applying the GPS to a new problem, but may cause trouble with the SARS, because the error behavior may be quite different, as was previously discussed. Fortunately, the user models do not share this problem. While better user trajectory modeling and estimation would undoubtedly improve the SARS' performance, there already exist proven kinematic user models that aid in accurate trajectory determination from GPS measurements [19]. Therefore, it seems likely that a large scale modeling effort is not warranted, but some thought does need to be given to the validity of the GPS error models used in the processing of the SARS data. As an example, most GPS receivers use a model of atmospheric delays to attempt to compensate for the effect of the atmosphere on the signals. If the standard GPS atmospheric models are used in the SARS, the receivers could over-compensate for the atmospheric errors actually present in the data, because in the SARS, the signals do not pass through the upper atmosphere or the ionosphere, where the majority of the distortion takes place. This overcompensation could diminish the system's accuracy. What this means is that the satellite-based GPS error models cannot be blindly transferred over to the SARS. The models need to be tailored to the SARS application.

1.2.2.4 Data Processing. The last main factor in GPS system accuracy is the way that the measurement data is processed to determine the user position or trajectory. There are only two kinds of measurements out of the GPS receiver (pseudorange and carrier beat frequency), but there are countless numbers of methods for using these measurements to determine the user's trajectory [3]. Although a detailed discussion of these methods is beyond the scope of this research, the basic concepts deserve to be mentioned.

The three basic areas where data processing affects GPS accuracy are the way the GPS measurements are used, the way the GPS measurements are filtered to remove errors, and the method used to solve the navigation equations. Each of these areas can have significant impact on the accuracy of the GPS system.

The basic methods of using the GPS measurements are called code-phase and carrier phase GPS. Code phase GPS uses only pseudorange measurements to solve for user position. Without differencing, this technique attains an accuracy window of about fifty feet. This is simply not good enough for the SARS. Fortunately, much work has been done on improved use of the measurement data to allow greater accuracy. The best current method for using the GPS measurements is called carrier phase GPS. In this method, the pseudoranges are used to obtain the fifty foot ballpark solution. Then, the carrier beat frequency measurement is used to estimate the total phase of the carrier wavefront from the receiver to the transmitter, locating the user position within the ballpark solution found using the pseudoranges [9]. This method can produce accuracies on the centimeter level, provided the estimator maintains 'lock' on the carrier phase. Obviously, given its potential high accuracy, carrier phase GPS will be used in the SARS, and has shown great promise in the initial testing at Holloman [17] and in follow on research at AFIT [2, 7].

There are several methods used to filter the errors out of the GPS measurements. Two standard techniques are error state estimation and differencing techniques. The error estimation techniques use models for the GPS errors to estimate the errors through Kalman filtering, and then remove the estimated errors from the measurements [11–13]. This kind of technique is very powerful given good models, but as was stated earlier, poor modeling can be worse than none at all. The models used had better be accurate. The differencing techniques involve taking measurements from adjacent receivers and transmitters, and differencing them to remove common errors. This kind of technique has proven effective in reducing GPS errors, and has been used in the concept testing of the SARS [17]. In most GPS applications, sophisticated differencing techniques cannot be used because of the need for real-time navigation capability. Fortunately, since the SARS data will be post-processed anyway for greater accuracy, differencing techniques for error removal can and will be used. For a complete description of carrier phase GPS differencing techniques evaluated for SARS use, see the works of Bohenek [2], Hebert [7], and Raquet, et al [17].

The final area where data processing affects the GPS accuracy is the method used in solving the non-linear navigation equations. In the standard GPS receivers, the navigation equations are solved using linearized equations and an iterative approach that converges on the solution. This iterative approach has proven successful for terrestrial navigation using the GPS satellites, but additional accuracy in regions of poor geometry (like the SARS) may be obtained by using a closed-form solution to the GPS navigation equations, as outlined in Bancroft's paper on the subject [1]. In fact, the closed form solution has already been used in GPS testing on ground based systems [18]. There, it was used to initialize the navigation and error estimation filters, as the closed form solution requires no initial a priori position information. This made the closed form solution an independent measurement suitable for updating the navigation filters. For the SARS, this closed form solution may be useful for two reasons. The first reason is that the closed form solution avoids convergence problems the iterative solution has in areas of poor geometry. The second reason is that the closed form solution may prove somewhat less sensitive to configuration geometry, obtaining greater accuracy in regions of high GDOP [1].

1.2.3 Promising Directions of Research. Now that the factors that affect GPS accuracy have been discussed, promising directions of research into improving GPS accuracy can be explored. The discussion of GPS accuracy suggests three possible areas of improvement: optimizing the GPS receiver array geometry, improving the SARS' GPS error models, and using the closed form solution to the navigation equations. Additionally, my exposure to navigation systems here at AFIT has suggested another possibility: aiding the SARS system with an additional (non GPS) sensor, such as a radar altimeter.

1.2.3.1 Optimizing the SARS Geometry. As discussed previously, the configuration of the reference stations and the user will impact the system's sensitivity to GPS errors. Therefore, it is desirable to optimize the placement of the reference stations on the ground (and the user flight profile) to minimize this sensitivity. Unfortunately, there does not seem to be any hard and fast design rule to use in such an optimization. What little information on this subject that appears in the literature has pertained to the satellites, and has not proven useful in ground based array (and user flight profile) optimization. Since the geometry will have such an impact on the SARS, and not much information on optimizing the geometry appears in the literature, it follows that research along this direction could prove quite useful in improving the accuracy of the SARS, by improving the choice of reference station locations on the ground and also the test flight profile. This optimization of geometry seems to be the most promising direction of research.

1.2.3.2 GPS Error Models. Another promising direction of research is refining the models used in the SARS. As previously discussed, there exist kinematic GPS models that can be used with carrier phase GPS to accurately determine the user trajectory. However, the validity of the models used for the GPS errors in the SARS data needs to be examined. If it indeed turns out that some of the error models used in the SARS right now are not valid, research should be done to refine them. From the discussion of the GPS error models, it looks like modeling improvements could be obtained in several ways: eliminating the ionospheric correction, modifying the atmospheric delay models to account for reduced signal path length, investigating the multipath errors, accounting for the larger noise to pseudorange ratio, and perhaps tuning the receiver models more 'tightly,' as all the receivers will be stationary. This modeling effort could yield some improvements in accuracy, but is likely to be of secondary importance compared to the SARS' geometry. Therefore, while refining the SARS models should be done to ensure highest accuracy, the research time seems better spent on the geometry optimization.

1.2.3.3 Use the Closed Form Navigation Equations. The third promising direction of research is investigating the benefits of using the closed form navigation equations. Bancroft states that using the nonlinear closed form solution has shown to be more accurate than the usual linear iterative technique used in most GPS algorithms, in regions of high GDOP [1]. This higher accuracy could make the closed form solution worth trying out in the SARS. Since all the data is post-processed anyway, trying out the closed form solution would be a simple matter of replacing a few lines of code in the program that executes the navigation equations. If increased accuracy around the poor GDOP regions of the SARS could be obtained, the useful area of the test range for flight testing purposes could be extended, and the flight profiles used would be less constrained. As converting to the closed form solution is such a simple change to the algorithm, it does not make sense not to try it out.

1.2.3.4 Use an Additional Sensor. If all else fails, a traditional method could be used to increase the system accuracy: aid the SARS with an additional (non GPS) sensor. It is desired to avoid this method, as it increases system complexity and adds to the hardware needed onboard the test aircraft. If nothing else has enabled the SARS to achieve its .1 meter or better accuracy goal, then aiding the SARS with an additional sensor would be worth investigating.

If this method is required, the problem then becomes deciding what kind of sensor(s) should be used. The carrier phase GPS measurements used in the SARS are highly accurate, much more so than most other sensors. Theoretically, any additional (accurate) information provided to the navigation solution should result in improved accuracy of the solution, but if the additional information is much less accurate, any improvements would be negligible. If properly modeled, the use of a sensor much less accurate than carrier phase GPS could still improve the SARS' accuracy, by reducing the GDOP. In this case, the addition of errors caused by the inaccurate sensor might be outweighed by the reduction in total error amplification gained by reducing the GDOP. Unfortunately, the chances of introducing error into the system would be significant, no matter what additional sensor is used, so an extensive validation effort would be needed to see if indeed any improvements were realized by the introduction of the sensor. Therefore, aiding the SARS with a non GPS sensor will not be investigated unless absolutely necessary.

1.2.4 Conclusion of Literature Review. This review has helped to obtain a clearer view of the research objective of improving the accuracy of GPS measurements at the SARS by identifying the factors that affect GPS accuracy and exploring possible directions for the research effort. The literature suggests that there are four main factors which affect GPS accuracy: GPS system errors, GPS system geometry, modeling, and data processing. Promising directions of research in these areas are optimizing the GPS system geometry. refining the SARS' error models, and using the closed form solution to the pseudorange equations. Out of these, the most promising looks to be optimizing the SARS' geometry. This thesis will concentrate on optimizing the SARS' geometry, leaving other possible improvements as topics for future research.

1.3 Problem Definition

The goal of this thesis is to improve the accuracy and usefulness of the SARS by optimizing the receiver locations to minimize geometric sensitivity. In the process of solving this problem, the geometric sensitivity of ground based arrays will be investigated and receiver array optimization tools will be developed. The knowledge and techniques from this preliminary research will allow the receiver array optimization problem to be solved and will provide CIGTF with a systematic method with which to optimize the SARS with respect to measurement geometry, allowing the SARS to be reconfigured whenever it is required.

1.4 Scope

Although there are many factors that affect the accuracy of an inverted GPS system like the SARS, only the geometry induced error sensitivity of the solutions to the pseudorange equations is addressed in this research. Such concerns as carrier phase ambiguity estimation, modelling of GPS errors, and data processing will not be addressed.

The work will be limited to:

1. The analysis and comparison of several measures of geometric sensitivity.
2. The development of MATLAB [10] programs to aid in the optimization and evaluation of inverted GPS receiver arrays.
3. The application of the developed software to the SARS' receiver array optimization problem.
4. The synthesis of the above research into a list of guidelines for good receiver placement for ground based receiver arrays.

1.5 Assumptions

This list shows the assumptions that were made to allow the research to focus in on the system *geometry*, as opposed to the many other considerations involved with carrier phase GPS.

1. Since this research is mainly geometry oriented, it is assumed that no 'cycle slips' occur in the carrier phase GPS measurements. Cycle slips are errors in the estimate of the number of carrier wavelengths (cycles) between the receiver and the transmitter [9]. These errors do impact the accuracy of carrier phase GPS, but the design of methods to detect and compensate for cycle slips is beyond the scope of this research. Cycle slip detection and compensation

techniques are discussed in the work of Bohenek [2]. Failure detection and compensation for GPS systems in general are discussed in the works of Mosle [15] and Vasquez [23].

2. Line of sight (LOS) from the transmitter to the receivers is assumed through most of this research. The reason why is that the LOS check developed for this research in MATLAB is *extremely* slow; it simply takes too long to be of use. Unfortunately, since the terrain around CIGTF is mountainous, the assumption of unbroken LOS throughout the entire flight may be questionable. Therefore, the LOS check (as slow as it is) is used several times throughout the research to check this assumption.

3. GPS errors like multipath etc. are not modelled in this research. They are assumed to be independent of the array geometry.

4. The curvature of the earth is ignored. Over the 50-100 mile extent of the SARS, the curvature of the earth will not have a significant effect.

5. The terrain is not taken into account for the main body of this research. Although elevation data for White Sands Missile Range is available, much of the research may be more easily done using a completely flat ground plane. It is felt that the development of the receiver array optimization tools will be more straightforward if the ground is assumed to be completely flat. Then, once the tools are developed, the elevation data will be used to see how much the altitude variations in terrain change the results.

6. There are no restrictions on where the receivers can go, except the boundaries of the test range. Although there surely are places inside the range where the receivers cannot go, most likely such places are small enough in extent that they can safely be ignored.

7. The range limits of the transmitter-receiver link are not incorporated into this research. The GPS transmitters and receivers used in the SARS are being

custom made for this application. It is assumed that they will prove adequate for the task. If not, a penalty on exceeding the range limits can easily be incorporated into the receiver array design tools.

1.6 Methodology

The two main steps to this research are shown in sections 1.6.1 and 1.6.2. Step one must be completed before step two can be begun. Although the results of step two will most likely benefit CIGTF the most, the results of both steps will be equally important to this thesis.

1.6.1 Step 1: Preliminary Research. Before the receiver array geometry optimization problem can be solved, the geometric sensitivity of the SARS must be analyzed and receiver array optimization tools developed. The analysis of geometric sensitivity is to better characterize the GDOP function and to explore alternate measures of geometric sensitivity that may be more meaningful than the GDOP. This analysis is the first order of business. Once it is done, the best measure of geometric sensitivity will be chosen for use in the optimization tools. The development of receiver array optimization tools is the next order of business. Three optimization tools will be developed and tested: a graphical technique for evaluating receiver array geometry, a Monte Carlo search program to help find the best receiver locations, and a systematic receiver array optimization program using **constr.m**: a Sequential Quadratic Programming (SQP) constrained optimization program from the MATLAB Optimization Toolbox. Once the right measure of geometric sensitivity is found, these three optimization tools will be developed and then put to work on the SARS' receiver array optimization problem.

1.6.2 Step 2: Receiver Array Optimization. The second and final step in the research is to use the knowledge and optimization tools developed in step one to solve the SARS' receiver array optimization problem to minimize the geometric sensitivity. This is to be done in three parts: the optimization of the receiver array with respect to a single, fixed transmitter location, the optimization of the receiver array with respect to multiple transmitter locations, and the testing of the assumptions 2 and 5 to see if they are indeed valid. The optimization with respect to a single, fixed transmitter location, although not an intended condition for the SARS, is included in this research because it can be simplified and solved in a way that the more general case of multiple transmitter locations cannot, providing valuable insights that are not readily apparent if only the multiple transmitter location problem is considered. Armed with the insights gained from the simpler problem, the case of multiple transmitter locations can be tackled using the optimization tools developed in step one. Finally, the results of the receiver array optimizations on flat ground will be checked on the White Sands Missile Range elevation data to see if the assumptions 2 and 5, terrain and LOS, are in fact reasonable or if they have an effect on the results.

1.7 Overview of Thesis

Chapter II presents the background material and theory used in this research. GPS technology is introduced to lay the groundwork for the understanding of the operation of the SARS and the importance of system geometry. Next, the SARS concept is discussed, examining similarities and differences between the SARS and satellite GPS. Finally, geometric sensitivity of GPS systems is examined to determine why it occurs and how it will impact the SARS.

Chapter III discusses the preliminary research mentioned in section 1.6.1. The geometric sensitivity of ground based receiver arrays is examined and alternate measures of geometric sensitivity are evaluated to find the best one for the purposes of this research. The SARS receiver array optimization problem is posed. Several optimization tools, a visualization tool and two receiver array optimization tools, are then developed and evaluated.

Chapter IV discusses the solutions to the receiver array optimization problems (single or multiple transmitter locations). The solution to the receiver array optimization problem for the single transmitter location is presented and array design insights are listed. The results of the more complicated array optimization problem for multiple transmitter locations are then discussed.

In Chapter V, all the results are summarized, a list of key design issues are given to aid CIGTF in the placement of the SARS' receivers, and a technique is shown that may allow CIGTF to significantly improve the geometry and therefore accuracy of the SARS beyond the capabilities of a totally ground based receiver array.

II. Background

This chapter presents the background material and theory used in this research. First, the basics of GPS technology are reviewed. Next, the SARS concept is discussed, examining the similarities and differences between the SARS and satellite GPS. Finally, geometric sensitivity is then examined, showing why it occurs and how it can be measured.

2.1 The Global Positioning System

The NAVSTAR Global Positioning System is a space-based radio-navigation system. It allows high position and time accuracy to be obtained with minimal user equipment. This system is used on many Air Force aircraft. This section reviews the basics of GPS, to aid in the understanding of the SARS concept and to help show the similarities and differences between the NAVSTAR GPS and the SARS. This background is by necessity brief. For further background into the workings of GPS, refer to the seminal article of Milliken, et. al. [14]. The Global Positioning System has three parts: the space segment, the control segment, and the user segment.

2.1.1 Space Segment.

The space segment consists of GPS transmitters mounted on satellites with carefully controlled, 12 hour orbits. The spacing of the satellites is arranged so that four or more satellites are in view at any given time, at any location. The satellites transmit time-tagged information back to Earth, where GPS receivers measure time difference of arrival of the signals. This requires the satellites to be synchronized. The satellites have highly accurate atomic clocks onboard, to make the time information as accurate as possible. The satellites transmit on two L-band frequencies called L_1 and L_2. L_1 is 1575.42 MHz and L_2 is 1227.6 MHz. L_1 is modulated by two pseudorandom codes: the C/A (coarse/acquisition) code and the P (precision) code. L_2 is modulated by just the P code. The C/A code has a frequency of 1.023 MHz, and is repeated every 1 millisecond. The P code has a frequency of 10.23 MHz, and is one week long. The C/A code is used to acquire the GPS signal; the P code is used once the C/A code is acquired to obtain higher accuracy from the system. The code sequences broadcast by each GPS satellite are different, so the GPS receiver can recognize each satellite.

2.1.2 Control Segment. The control segment consists of monitor stations and a master control station. The monitor stations, spread across the world, passively track the satellite signals to acquire accurate ranging data. This is used to pinpoint the satellite positions in space. This data is sent to the master control station, which uses it to correct the satellites' orbits and upload corrected navigation message data to the satellites. The control segment also maintains GPS time, perhaps the most important part of the system.

2.1.3 User Segment. The user segment consists of a GPS receiver that demodulates the data, and a processor that uses the received information to calculate the receiver's position and time. The receiver provides the processor with time difference of arrival measurements, called pseudoranges, from each satellite in view. These pseudoranges are used via multilateration to find the position of the GPS receiver. The computation of the receiver's position requires the solution of the pseudorange equations, shown here for the four receiver case. The pseudorange equations are:

$$\begin{aligned}
(x_1 - u_x)^2 + (y_1 - u_y)^2 + (z_1 - u_z)^2 &= (R_1 - B)^2 \\
(x_2 - u_x)^2 + (y_2 - u_y)^2 + (z_2 - u_z)^2 &= (R_2 - B)^2 \\
(x_3 - u_x)^2 + (y_3 - u_y)^2 + (z_3 - u_z)^2 &= (R_3 - B)^2 \\
(x_4 - u_x)^2 + (y_4 - u_y)^2 + (z_4 - u_z)^2 &= (R_4 - B)^2
\end{aligned}$$

where the pseudorange R_i to the satellite i is defined as $c\Delta t$ and:

- c = speed of light

- $\Delta t_{i=1,2,3,4}$ =signal transit time
- $(x_{i=1,2,3,4}, y_{i=1,2,3,4}, z_{i=1,2,3,4})$ are the respective i-th satellite positions
- (u_x, u_y, u_z) is the GPS receiver position that is being solved for
- B is the GPS receiver clock bias

To obtain a solution to these equations, at least four satellites need to be in view. The above example is for the four satellite case. More than four satellites can be used, if desired. In that case, least squares or some other estimation technique is used to solve the overdetermined set of equations. To obtain the highest accuracy, it is desired to have as many satellites in view as possible.

The pseudorange equations can be solved either as shown above, or they can be linearized to save computation time. Although an algebraic solution to the pseudorange equations may be desirable from the point of accuracy [4], most GPS receivers in use today use the linearized equations to save computation time. This allows the position fix to be updated at shorter intervals than if the algebraic solution is used [20]. The linearized navigation equations are:

$$\begin{bmatrix} \cos\theta_{1x} & \cos\theta_{1y} & \cos\theta_{1z} & 1 \\ \cos\theta_{2x} & \cos\theta_{2y} & \cos\theta_{2z} & 1 \\ \cos\theta_{3x} & \cos\theta_{3y} & \cos\theta_{3z} & 1 \\ \cos\theta_{4x} & \cos\theta_{4y} & \cos\theta_{4z} & 1 \end{bmatrix} \times \begin{bmatrix} \Delta u_x \\ \Delta u_y \\ \Delta u_z \\ \Delta B \end{bmatrix} = \begin{bmatrix} \Delta R_1 \\ \Delta R_2 \\ \Delta R_3 \\ \Delta R_4 \end{bmatrix}$$

These equations can be written more compactly as $H\mathbf{x} = \mathbf{r}$. $\mathbf{r}$ is the vector of estimated pseudorange errors, $\mathbf{x}$ is the vector of estimated user position and time errors, and H is a matrix of coefficients.

These linear equations represent the error behavior of the navigation equations about some nominal user (GPS receiver) position. Of course, these error equations are meaningless without an accurate nominal position. This is a problem, because the position of the GPS receiver is the unknown quantity that needs to be found. This problem is overcome by using functional iteration to converge upon good estimates of $\mathbf{r}$, H and $\mathbf{x}$. The technique works as follows:

1. **A nominal receiver position and time is selected.** For satellite GPS, the initial nominal GPS receiver position is chosen to be the center of the earth, and a zero clock bias is chosen.

2. **The H matrix is formed.** The positions of the GPS satellites are calculated by the GPS receiver from ephemeris data received in the GPS signals from the satellites. The H matrix is formed, given the positions of the GPS satellites and the nominal position of the receiver.

3. **The $\mathbf{r}$ vector is calculated.** The vector of pseudorange errors is calculated. This is the difference between the vector of pseudorange measurements and the vector of pseudoranges calculated using the nominal receiver position and the positions of the GPS satellites.

4. **The equations are solved for $\mathbf{x}$.** The system of equations is solved for the estimated receiver position errors $\mathbf{x}$ using least squares.

5. **The error is removed.** The estimated position (and time) errors are subtracted from the nominal position and time estimate. This produces an estimate that is closer to the actual GPS receiver position and clock bias than the previous estimate.

6. **Repeat until convergence.** This algorithm is repeated until the quantities H, $\mathbf{r}$, and $\mathbf{x}$ converge, i.e., don't change with successive iterations. Once converged, an accurate estimate of the GPS receiver position and clock bias is known.

Once the iterative technique has converged, the nominal GPS position and time estimates using C/A code are reasonably accurate, i.e., to within 30 meters. Additional accuracy is obtained by using differencing techniques, P code, and filtering techniques that estimate the carrier phase. Depending on the particular GPS application, any combination of these techniques may be used.

Differencing is where more than one receiver is used and the measurements are differenced to remove errors common to both receivers. Many errors in the pseudoranges can be reduced or eliminated through differencing techniques. Differencing takes the C/A code error down to within 10 meters. This technique is in common use in GPS navigation.

Additional accuracy can be obtained by using the P code. P code is at ten times the frequency of C/A code. Using P code provides ten times the accuracy of C/A code GPS. Since P code is transmitted on both L_1 and L_2, frequency dependent errors can be calibrated out of the system without use of a second reference receiver. As the P code is classified, it can only be used for certain applications (like on Air Force aircraft). Differential P code obtains accuracies of within one to two meters.

Estimating the phase of the carrier of the GPS signal can obtain additional accuracy. The most complicated of all the techniques discussed here, carrier phase estimation has the greatest potential accuracy. The GPS receiver has the capability to measure the phase of the carrier to within a small fraction of a wavelength [3]. The carrier wavelength is on the order of .2 meters. Therefore, the potential accuracy of this technique is on the centimeter level. However, there are some difficulties. The fractional part of the phase can be measured, but the rest of the total carrier phase (the integral number of wavelengths between each satellite and the receiver) must be estimated. It is difficult to estimate this number of wavelengths (called the carrier phase ambiguity). If anything interferes with the measurements (such as a change in satellites, vehicle dynamics, or loss of lock), the carrier phase ambiguity estimates may suddenly change. This is called a 'cycle slip.' The carrier phase estimation filter must be able to process not only GPS signal information at a high rate, but also simultaneously run a cycle slip detection and compensation routine [2]. It is difficult to process this information fast enough to provide real time carrier phase GPS capability in a highly dynamic environment. Therefore, for the GPS applications that require the highest accuracy, post processed carrier phase GPS is employed.

Although these techniques all improve the accuracy of the GPS position estimate, no technique can remove all the error. Some pseudorange error $\delta\mathbf{r}$ always remains after compensation. The residual position error is given by $H^{-1}\delta\mathbf{r}$, for the four satellite case. In the case of more than four satellites, least squares (in place of the inverse) is used to solve this equation. For a given level of residual pseudorange error, the amount of residual position error depends on H. The H matrix is composed of direction cosines between pseudorange vectors and the coordinate axes. The i-th row of H represents the direction cosines of the i-th pseudorange vector with each axis of the four dimensional (space-time) Cartesian coordinate system: (x,y,z,t). Since the H matrix is composed entirely of direction cosines, it is purely a function of GPS measurement geometry. Therefore, the measurement geometry of GPS systems has a profound impact on the accuracy of GPS position estimates. Fortunately, the GPS satellite constellation is designed to minimize the effect of geometry on GPS accuracy [14, 20]. Usually, geometry is not a problem with satellite GPS.

Summary:

The Global Positioning System provides a worldwide, high accuracy navigation system. It requires little user equipment and provides high accuracy in position and time. It maintains accuracy over long periods of time and during a wide variety of conditions. The GPS receiver uses C/A code pseudorange measurements and satellite position data to solve for its position within 30 meters. Depending on the application, the techniques of differencing, using P code, and estimating the carrier phase may be used to improve the accuracy of GPS beyond the base accuracy of C/A code GPS. The measurement geometry of GPS does affect its accuracy, but this effect is minimized by the design of the GPS satellite constellation. GPS provides worldwide, high accuracy position estimates, which makes it ideally suited for use onboard Air Force aircraft and has been integrated into the navigation systems of many. This has caused CIGTF difficulty, as it has proven difficult to accurately flight test these integrated navigation systems. Hence, CIGTF is in the process of developing the SARS, with its .1 meter accuracy, to enable accurate flight tests of these integrated navigation systems to be performed.

2.2 The Submeter Accuracy Reference System

The Submeter Accuracy Reference System is CIGTF's new navigation reference system under development. It uses carrier phase GPS technology and post processing to obtain high accuracy [7, 17]. This accuracy will enable the SARS to accurately test some of the better navigation systems on Air Force aircraft, something which has been a problem in the past.

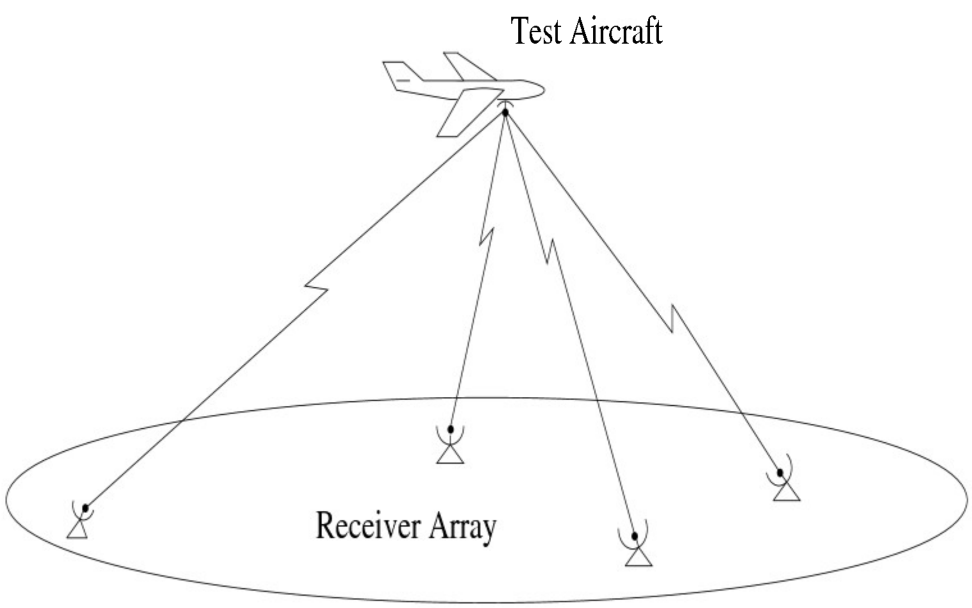


Figure 2.1 SARS Concept

The concept of the SARS is shown in Fig. 2.1. The basic concept of the SARS is to use an array of GPS receivers to locate the position of a moving GPS transmitter, called a pseudolite. This technique is independent of the GPS satellites, allowing performance of GPS-based aircraft navigation systems to be accurately tested. Since the transmitter and receivers are both controlled by CIGTF, they can be made to operate at different frequencies than the satellites, allowing the reference system to test aircraft navigation system performance under jamming. Most importantly, this system requires only a transmitter to be mounted on the test aircraft. This allows the system to test the navigation systems of all Air Force aircraft.

The SARS uses carrier phase GPS to obtain the high accuracy it requires. This requires a reference pseudolite and an array of GPS receivers at very accurately surveyed locations, in addition to the mobile pseudolite [7,17]. The mobile pseudolite is a GPS transmitter attached to the test aircraft. During the flight test, each GPS receiver in the receiver array records the pseudoranges to both the stationary and mobile transmitters. After the flight, this information is transferred to the central processing station. There, the raw data is processed and the carrier phase estimation filters are run. The end result is an accurate recording of the aircraft's trajectory during the flight test. This is used to evaluate the performance of the aircraft's navigation system. For a detailed discussion of carrier phase estimation as applied to the SARS, refer to the work of Hebert [7].

Although the SARS is not yet fully operational, the concept has been tested [17]. The concept test used a GPS transmitter mounted on a van as the mobile pseudolite. The receiver array consisted of six GPS receivers. Although the geometry of the configuration was poor because everything was nearly coplanar, the system performed well. Accuracies of 10 to 30 centimeters were obtained. This is encouraging news for the SARS, because it shows that only a small improvement in accuracy is needed to obtain the desired accuracy goal of .1 meter. As the geometry for the concept test was so poor, it is anticipated that optimizing the geometry of the SARS will achieve the desired improvement in accuracy [7].

As it stands now, the SARS' concept has been successfully tested, but its accuracy needs to be improved as much as possible to enable it to perform its mission.

It is thought that the SARS' measurement geometry will profoundly affect its accuracy. Unfortunately, the best ways to place the SARS' receivers on the ground are not known. The receivers will be located on the ground, nearly in a flat plane. This kind of planar configuration is avoided in the design of the satellite GPS. The literature on satellite GPS geometry does not aid in the design of SARS receiver arrays. Therefore, the geometric sensitivity of ground based GPS systems such as the SARS need to be examined to enable intelligent choices of receiver array configurations to be made.

2.3 Geometric Sensitivity

Geometric sensitivity is due to the fact that pseudorange errors are magnified in the solution of the pseudorange equations for position. Geometric sensitivity is a gauge of the goodness of the position estimate. The degree of error magnification depends on the relative positions of the GPS system elements (transmitter and receivers, for the SARS). The information about the system measurement geometry is contained in the H matrix, also called the *visibility* matrix. The H matrix is composed of angle cosines; each row contains the cosines of the angles between the line of sight vector from the transmitter to receiver and each of the component coordinate axes: $(\cos x, \cos y, \cos z, \cos t)$. The cosines with respect to the time dimension are always 1 for every receiver. Although the H matrix is composed of direction cosines, it is somewhat different from the familiar direction cosine matrix. A direction cosine matrix represents a coordinate transformation between orthonormal bases (and is therefore itself orthonormal). The H matrix can be considered a *non-orthogonal* transformation from a basis consisting of four or more transmitter-receiver line-of-sight vectors to the four dimensional space consisting of an orthogonal Cartesian coordinate basis and the time dimension. Therefore, the H matrix does not have the desirable property of orthonormality, as does the direction cosine matrix. This can be a problem, because the solution process involves inverting the H matrix. If the matrix is ill conditioned, small errors in the measurements (the 'inputs') can have a big impact on the computed position, (the 'output'), as shown by the following 'contrived' example.

Consider the equation $Ax = z$.

$$\begin{bmatrix} 1 & 1000 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 1000 \\ 1 \end{bmatrix} \tag{2.1}$$

Solving the matrix equation for a and b yields $a = 0$ and $b = 1$. Consider now the same equation, but with a small error δz in the input z.

$$\begin{bmatrix} 1 & 1000 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 1000 \\ 1 + \delta z \end{bmatrix} \tag{2.2}$$

Solving the equation now yields $a = -1000\delta z$ and $b = 1 + \delta z$. The relative error in the input vector z is $\frac{\delta z}{1000}$. The relative error in the output vector x is approximately $1000\delta z$. Therefore, solving the system of equations amplifies the relative input/output error by a factor of 10^6. Of course, this is a contrived example, but the same kind of thing can happen with the GPS pseudorange equations.

Indeed, it is possible for the H matrix to become ill conditioned as a result of the system geometric configuration. The receiver array and transmitter configurations that result in an ill conditioned or nearly singular H matrix are generally referred to as 'poor geometry.' Examples of good and bad GPS measurement geometry are shown in Fig. 2.2. One could run into trouble due to an ill conditioned or nearly singular H matrix (poor geometry) in three ways. The first way is when the receiver array and the transmitter all lie approximately in one plane. This results in a nearly linearly dependent column in H. It does not matter how many receivers there are; if the transmitter is nearly coplanar to them, the geometry will be very poor and the errors in the position will be greatly amplified. Therefore, the receiver array should

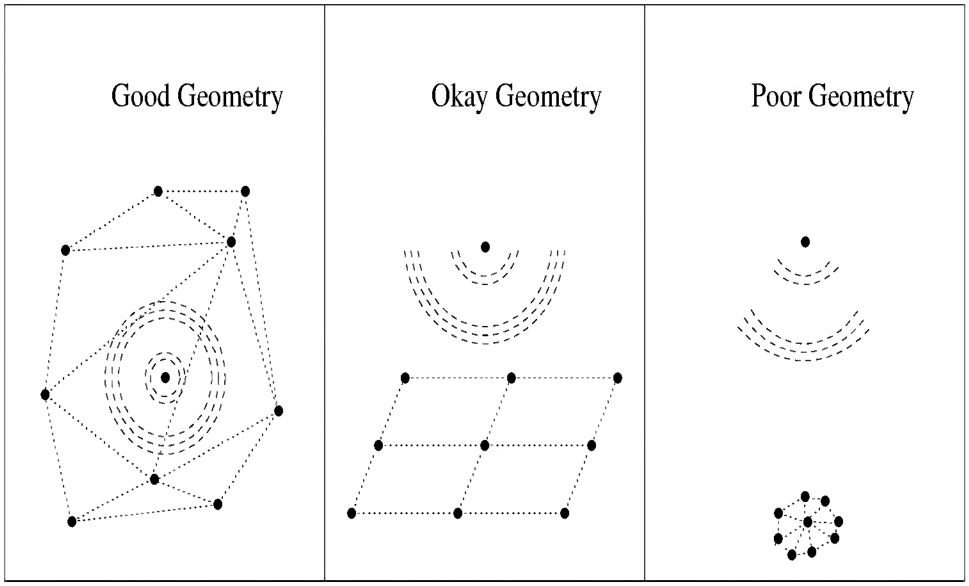


Figure 2.2 4 Receiver Array, Optimized about the Origin

be designed so that either all the receivers do not lie in one plane or the transmitter will never need to approach the plane of the receiver array. The second way to get poor geometry is when the line of sight vectors drawn from the transmitter to the receivers have similar angular orientations. This occurs when several receivers 'line up' from the viewpoint of the transmitter, as happens when the transmitter is very far from the receiver array. This idea can be visualized as an eclipse: receivers appearing to be in the same part of the sky from the viewpoint of the transmitter. The eclipse analogy is particularly apt: the 'eclipsed' receivers do not have any impact on the geometry. They could be removed from the array without changing the geometry at all. This occurs because the H matrix will have nearly equal rows for those receivers that are eclipsed. Only one of those nearly linearly dependent rows will have any geometric significance. Therefore, one must either have enough receivers to withstand 'eclipses' or design the array as to avoid them. The third way to run into trouble is for the transmitter to be in certain 'unfavorable' positions near a receiver array where the solution does not exist. An example of such an unfavorable location is when the transmitter is placed directly over the center of a planar array composed of a four receiver square. This produces a singular H matrix, signifying that the solution does not exist for that configuration. Clearly, receiver arrays with a high degree of symmetry that produce such 'unfavorable' locations must be avoided as much as possible.

Summary:

The geometric sensitivity of the GPS position estimate is shown in the H matrix. The H matrix is a function of the relative positions of the GPS receivers and transmitter (for inverted GPS). The H matrix is the mapping between residual pseudorange errors and errors in the position estimate. Since all the pseudorange error cannot be eliminated, there always is some position error. For a given amount of residual pseudorange error, the amount of position error is determined by H. The GPS measurement configurations that produce little amplification of pseudorange errors due to H are called 'good geometry.' Although it is known that the GPS satellite constellation produces good geometry for satellite GPS, it is not known what receiver array configurations will produce good geometry for the SARS. It is desired to find good receiver array configurations for the SARS so that its accuracy can be improved as much as possible.

III. Preliminary Research

This chapter discusses the preliminary research into the receiver array optimization problem. Preliminary research was needed in order to properly set up the receiver array optimization problem and develop numerical optimization tools to aid in its solution. The preliminary research consists of the following: the characterization of GDOP, an examination of alternate measures of geometric sensitivity, the proper choice of cost function in the receiver array optimizations, the posing of the receiver array optimization problem, and the development of receiver array optimization tools to aid in its solution.

3.1 The Characterization of GDOP

Throughout the literature, the GDOP is used as a measure of geometric sensitivity for GPS applications. The behavior of GDOP on the ground for satellite GPS is fairly well documented, but SARS' geometry is different enough from satellite GPS to warrant an examination of the behavior of GDOP for the SARS. The examination of the GDOP is conducted in two parts. First, a simplified two dimensional (planar) GPS navigation example is discussed to gain insights into the more difficult three dimensional case. Next, the analysis of GDOP for three dimensional navigation is discussed. The results of this analysis reveal the behavior of GDOP as a function of receiver array configuration and transmitter position. These results show how the SARS geometry is so much different than that of the satellite GPS and why the SARS' receiver arrays need to be carefully designed if high accuracy is to be obtained.

3.1.1 Simplified (2D) GDOP Analysis. GDOP is a difficult function to minimize. The GDOP function is the square root of the trace of a matrix inverse, which makes it difficult to minimize analytically. The matrix being inverted (H) is the matrix that maps the pseudorange measurement errors into position errors.

Its complexity increases with the number of simultaneous measurements (number of GPS receivers). It is desired to get a 'handle' on the behavior of GDOP around GPS receiver arrays (for inverted GPS), to gain insights that will aid in the optimization of the receiver arrays.

The first step in analyzing the GDOP is to consider the simplified case of two dimensional navigation. The two dimensional case removes the third dimension from the picture. What is left is a planar navigation problem, which requires three GPS measurements (or more) in order to solve for the three errors (two position errors, one transmitter clock error). The problem then becomes determining the planar region(s) of acceptable GDOP around the two dimensional receiver arrays. The analysis of this model provides insight into the general way GDOP depends on the system configuration, and gives a rough idea of the number of receivers and general configuration needed to obtain a region of good GDOP large enough to conduct a reasonable flight test within.

3.1.1.1 MATLAB 2D Range Simulation. Several MATLAB programs are used to take a specified receiver array and calculate the GDOP at the transmitter as it flew by on some simple trajectory (line, circle, etc.). These algorithms are designed to provide plots of the general behavior of GDOP as the receiver array design parameters of receiver configuration, receiver spacing, baseline, number of receivers, and transmitter position are varied. The use of these algorithms helps to identify the behavior of GDOP and to give a rough estimate for how big a viable range a certain configuration or number of receivers produces. The programs discussed are simple and general. Receiver and transmitter positions are given as simple points on the x-y plane. Receiver spacing, although it does differ among the configurations, was set at approximately 1 unit to simplify the analysis. This is justified because GDOP is a function of the geometry (shape), but not a function of the size of the shape. This simplifies the model analysis because range from transmitter to the receivers can be normalized to the spacing between the receivers (or the receiver baseline). There are three different transmitter flight profiles considered: a circle around the center of the receiver configuration, a line of constant distance from the receiver baseline, and a line moving directly towards/away from the array of receivers. These three flight profiles allow the two dimensional simulation to be considered to lie within a horizontal plane or a vertical plane. Since the algorithm takes only relative positioning into account, this interpretation is justified. (One could call the position axis altitude,westward position, or some other axis orientation without loss of generality.)

3.1.1.2 2D Simulation Results. The Matlab simulations provide the expected insight into the behavior of the GDOP as a function of the system configuration. Many different receiver configurations and transmitter flight profiles are considered, allowing a full characterization of the GDOP function. The results of the simulation indicate several overall trends that prove helpful in the design process. The first trend is that GDOP seems to depend mainly on the ratio between the receiver baseline (receiver array length) and the distance between the transmitter and the receiver array. The second is that most receiver configurations produce regions of alternating good and terrible GDOP as the transmitter moves along its flight profile. Those arrays are not acceptable for use in the test range, regardless of how low the minimum value of GDOP in the good parts is. The last trend is that the overall system performance improves gradually with the addition of more GPS receivers, with the point of diminishing returns occurring at such a large number of receivers that other constraints will ultimately determine the number of receivers actually used in the system.

The first trend is the most obvious. GDOP is mainly a function of the ratio between the distance from the transmitter to the center of the receiver baseline and the length of the receiver baseline. In fact, if the receiver baseline were to be defined as the width of the receiver array perpendicular to the line of sight between the transmitter and the center of the array, GDOP would almost entirely be a function of that transmitter distance/receiver baseline ratio. It turns out that this result is a relation between GDOP and the angle that the receiver array occupies from the transmitter's point of view. The relation between GDOP and this angle, called the field of view (FOV), is investigated later on in this chapter. An example of the relation between GDOP and the FOV in two dimensions is shown in Fig. 3.1.

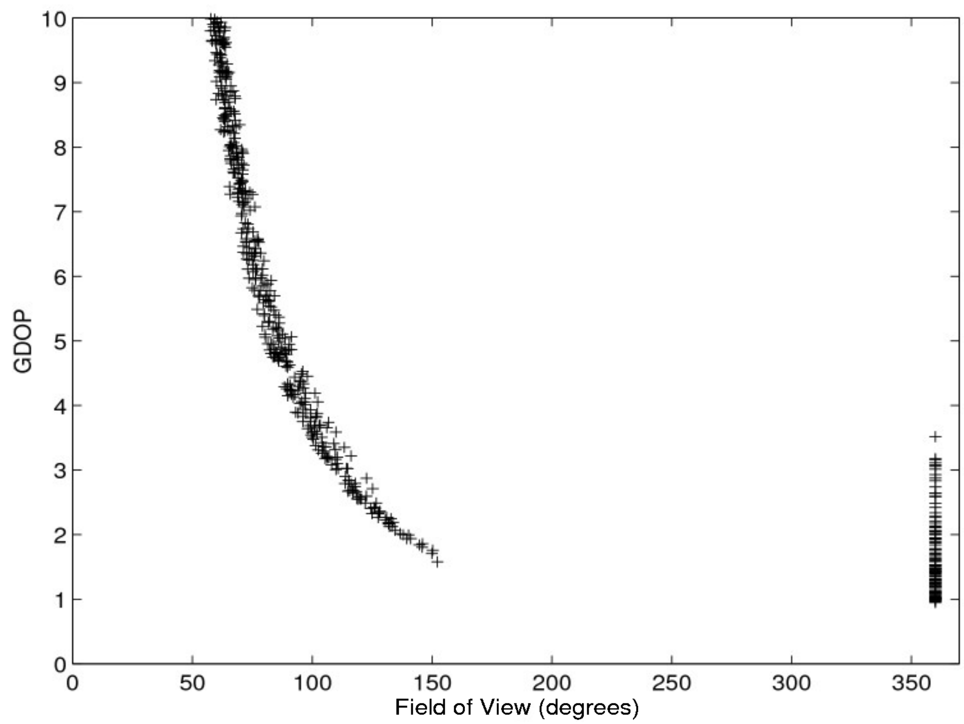


Figure 3.1 GDOP vs. Field of View

The GDOP quickly gets large as the transmitter distance/array size ratio increases. It may also get large as the ratio decreases past a certain value, depending on the receiver array. An example of this is shown in Fig. 3.2. If the array is a line of receivers, as is shown in Fig. 3.2, the GDOP will be large for both extremes of the ratio. If the array deviates significantly from a linear shape, the GDOP will only get large when the ratio gets large. It has been found that the GDOP with the transmitter inside such an array is better than the GDOP with the transmitter outside the receiver configuration. Unfortunately, this is not really possible with a

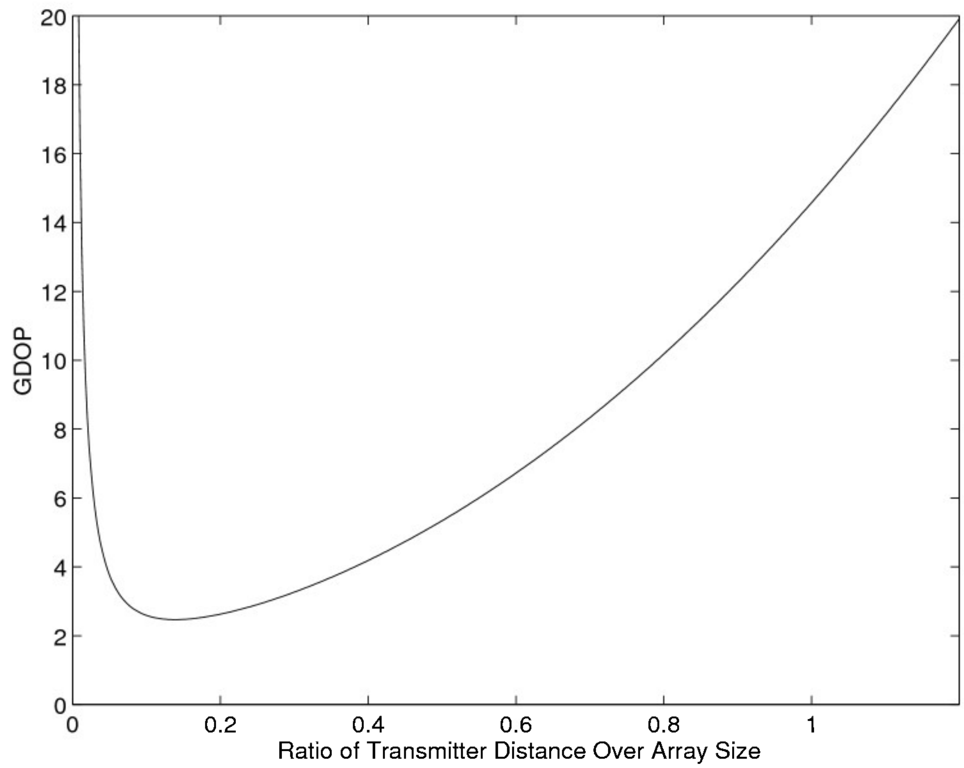


Figure 3.2 GDOP vs. Ratio of Transmitter Distance to Array Size

ground based receiver array. (Note: when the receiver array is nearly collinear, this does not work, and turns out to be one of the worst places to put the transmitter). This trend affects the system design in that it limits the altitude range over which any given configuration will have good GDOP.

Good GDOP is attained only in a narrow altitude band directly over the receiver array. Both minimum and maximum altitude bounds exist, for any desired GDOP. The minimum bound occurs because lines drawn from the transmitter to each GPS receiver in the array are almost collinear at low altitudes. This reduces the effective number of receivers in the array to just two, making the GDOP in certain positions skyrocket. If the transmitter is directly above a GPS receiver, this problem does not occur. The effective number of GPS receivers will still be three, which is enough to solve the GPS equations. With a fixed, ground based receiver array, it is not possible to keep the transmitter directly above a GPS receiver at all times during the flight. A minimum altitude restriction is therefore imposed on the aircraft, preventing it from flying into the bad GDOP regions close to the ground, in between the receivers. Using a moving receiver or a GPS satellite could eliminate this minimum altitude problem, if necessary. The maximum altitude bound occurs because lines drawn from the transmitter to the receivers all seem to be going the same way if the transmitter is at high altitudes. This reduces the number of effective receivers in the array down to one receiver, clearly making the GDOP greatly increase. Therefore, for the two dimensional equivalent of the SARS, *there is a minimum altitude bound and a maximum altitude bound on the transmitter's flight profile.*

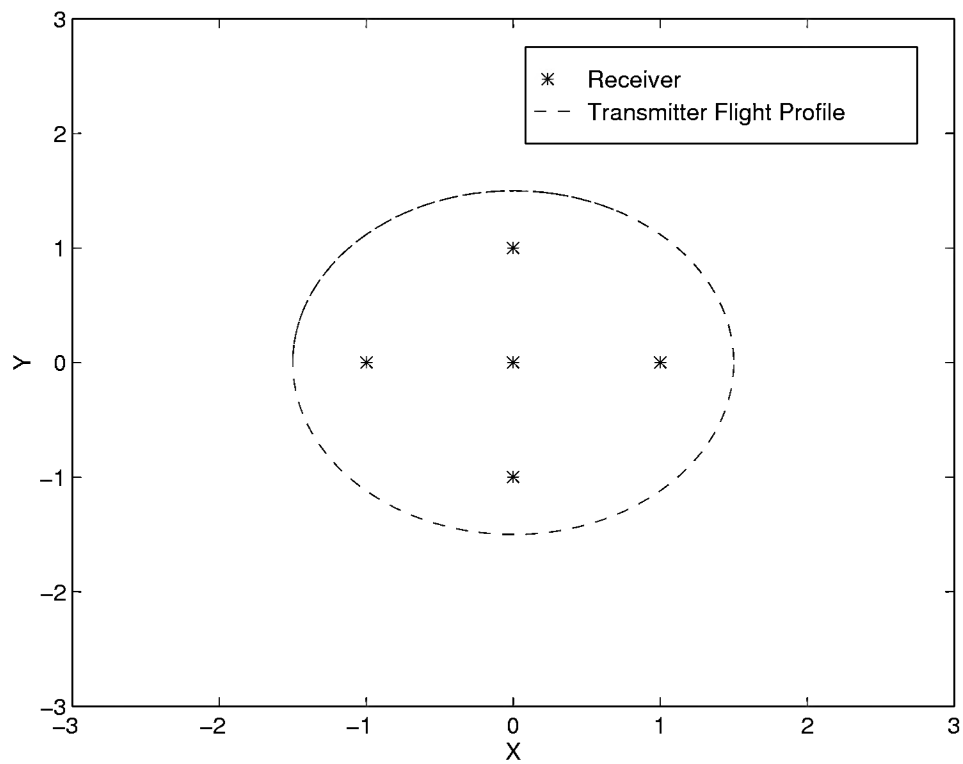


Figure 3.3   Simple Circle 'Flight' Around a Receiver Array

The second trend is that the GDOP is choppy for some receiver configurations. This was discovered by simulating the transmitter flying a circle around the receiver array, as shown in Figs. 3.3 and 3.4. This is easiest to interpret (in 3D) as flying

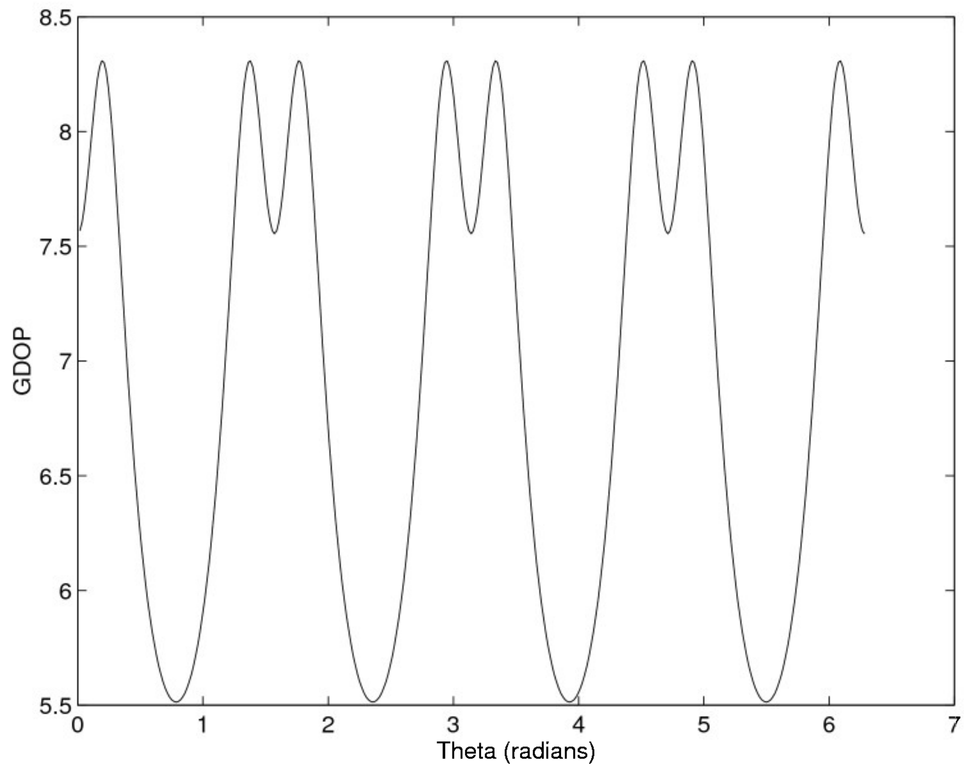


Figure 3.4 GDOP vs. Theta: Circle Profile in 2D

a circle of constant altitude and radius around the edge of a receiver array, but can also be seen as flying over the array and plotting the GDOP vs. transmitter elevation angle from the center of the receiver array. However one visualizes it, the simulation shows that the GDOP goes through sizeable maxima and minima as a function of the position of the transmitter on the circular flight path. This looks somewhat similar to the maxima and minima of a phased array antenna. It turns out that any time that a line can be drawn from the transmitter through two or more of the receivers, the GDOP sharply increases. That can possibly be interpreted as an effective reduction of one or more GPS receiver from the array, at that instant. This kind of trend points to making the receiver configuration be a flat plane (or a straight line in two dimensions) to maximize the angular region of good GDOP. Fortunately, that kind of configuration is what is natural for a ground based system, so this does not pose that big of a problem in the design. The biggest benefit of the planar (or straight line in 2D) configuration is that the receiver baseline is maximized (for when the transmitter is directly above the array), giving a large angular field of good GDOP directly overhead (inside the good altitude range).

The final trend is that increasing the number of receivers in the array gradually but continually improves the performance. The performance improvements bought by adding receivers are increased array baseline (by putting more on the end), larger region of low GDOP (mainly increases altitude range), overall lower GDOP, and smaller variation in GDOP within the good GDOP region. Increasing the receiver baseline by extending the array simply widens the region of good GDOP. The minimum altitude is not affected, but the maximum altitude is increased. The horizontal region of good GDOP is extended by the amount that the baseline is extended. One might think that simply increasing the receiver spacing would have the same effect, but it does not, because the minimum altitude increases as well. It has been found that sizable performance improvements are realized when receiver numbers are increased to about fifteen or twenty in a line, at which time a point of diminishing returns is reached. Of course, that's probably too many (especially if a grid is intended), so a tradeoff is necessary. The number of receivers in a line to get a GDOP of two for some small altitude range over most of the baseline is six or seven. The corresponding (small) good altitude range over that baseline is about three quarters of the distance between two consecutive receivers, starting at a minimum altitude of half that distance. That's a pretty tight region, so it may be worth it to have a large number of receivers in the array.

3.1.1.3 Summary: Simplified (2D) GDOP Analysis. The results indicate that there is a tradeoff between system utility and the number of receivers used in the array. With too few receivers in the array (approximately less than a 5×5 grid), the GDOP of the system is higher than that obtained from using the satellites. The baseline is too small, or the receivers are too widely spaced. If the baseline is too small, the test range is simply too small to be useful. If the receivers are too widely spaced, the minimum altitude is too high for some of the Air Forces aircraft or for any kind of low altitude mission. Without an external aiding measurement of some kind to improve the geometry, the only way to improve the system is to keep adding more receivers, which quickly gets complicated and costly. An array of many receivers would work, but would be time consuming to maintain, complicated to coordinate, and costly to acquire and upgrade. How many is too many? It really depends on the mission of the range, and the logistics of the situation. If low altitude, wide area testing needs to be done, then lots of receivers need to be used. If logistics are the prime concern, then maybe a high altitude, reduced area test range needs to be considered, if indeed it is worth it to implement it at all. The overall conclusion of this research into the two dimensional navigation problem is that unless lots of receivers are available, either a GPS satellite, a moving receiver, or an additional measurement (e.g., altitude) needs to be incorporated into the system to obtain the desired accuracy and size of the range necessary for this system to be useful. As a side note, the addition of just one GPS satellite to aircraft measurement so dramatically improves every aspect of the system performance that some method for using a GPS satellite (even during jamming) should be considered, if possible.

3.1.2 Three Dimensional GDOP Analysis. Now that a basic idea of what to expect has been provided by the two dimensional problem, the behavior of GDOP for three dimensional GPS navigation is examined. This is done by using graphical technique to plot the GDOP as a function of transmitter location for a fixed receiver array. The GDOP around a receiver array is plotted as a 3-D scalar field, much like a plot of voltage around charged particles. The positions of the receivers are known, and the GDOP at any point in space is taken as if the transmitter were to occupy that same point. In this way, the resulting three dimensional scalar field of GDOP indicates what the GDOP would be at any point in that space if the transmitter were located there. This provides three dimensional plots of GDOP 'fields' around receiver arrays. The GDOP fields around receiver arrays are plotted to help figure out which receiver array configurations produced the best overall GDOP characteristics. This entails a reasonable low GDOP range which is uniform, i.e., without 'spikes' of high GDOP. This plotting of the GDOP fields is very useful for receiver array design.

The plotting of the GDOP fields around different receiver arrays sheds light on the receiver array design problem. Examination of the plots reveals that there is considerable leeway in deciding what the 'best' GDOP field is, that the best receiver array configuration depends on the specific use of the array, and that the use of a ground based array places severe constraints upon the system design. These findings help clarify the design problem for the SARS and give important insights towards what could be done in the subsequent numerical optimization of the receiver array.

There is some leeway in deciding just what a good GDOP field is. Given some receiver array configuration, the GDOP field produced by the receiver array will be 'good' if the GDOP of the measurements is satisfactory over the aircraft's flight profile(s). Satisfactory GDOP certainly depends on the application. For the SARS, it is desired to have GDOP similar to that of the satellite GPS system, i.e. a GDOP less than 5 throughout as much of the flight profile as possible. Any number of array configurations could be made to provide the satisfactory GDOP for a given flight profile, with the same or different numbers of receivers in the arrays. As an example, Figs. 3.5-3.8 show two planar arrays that are different but provide essentially the same GDOP throughout the aircraft's flight profile. Similarly, if the receiver array configuration is fixed, the flight profile can be specifically chosen to make the most efficient use of the receiver array's GDOP fields, as shown in Figs. 3.9 and 3.10. If the resulting arrays and flight profiles produce satisfactory GDOP histories, the GDOP fields are classified as 'good'.

Of course, both of these methods require either the flight profile or the receiver array to be known in advance, before the 'goodness' of the GDOP fields can be evaluated.

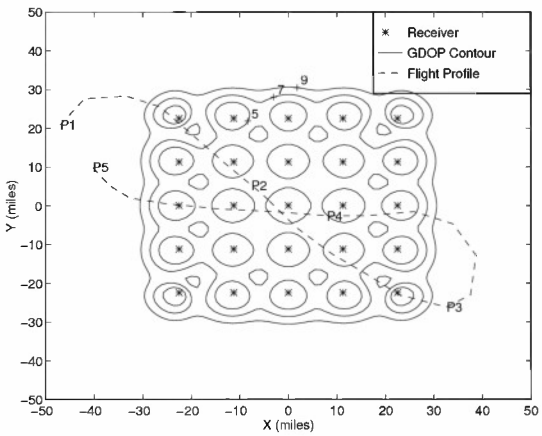


Figure 3.5 GDOP Fields, 1 mi. alt, Array 1

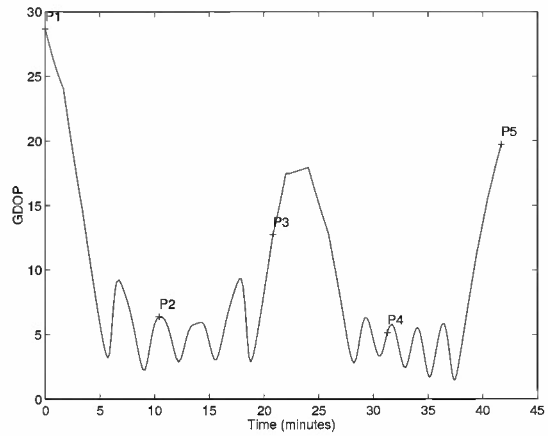


Figure 3.6 GDOP Along the Flight, Array 1

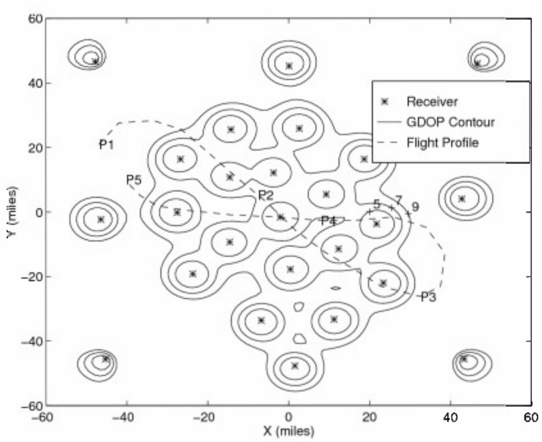


Figure 3.7 GDOP Fields, 1 mi. alt, Array 2

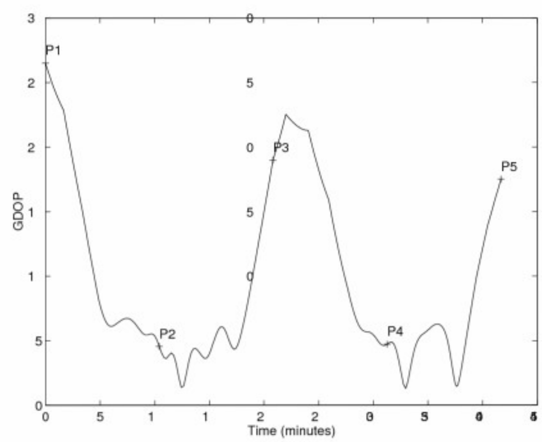


Figure 3.8 GDOP Along the Flight, Array 2

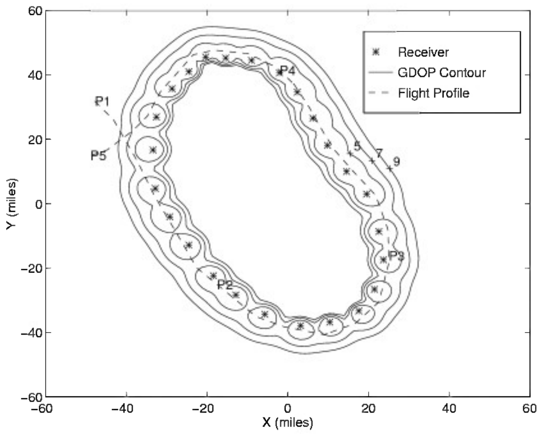


Figure 3.9 GDOP Fields, 1 mi. alt, Array 3

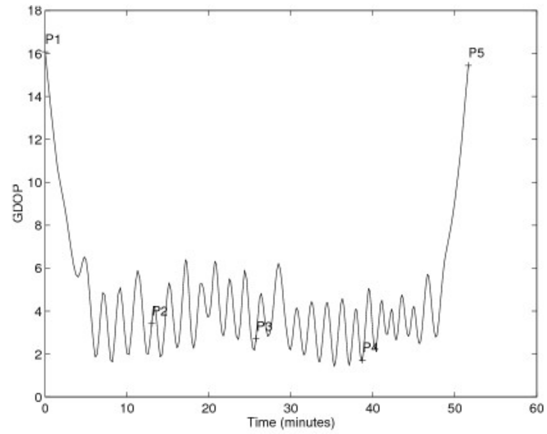


Figure 3.10 GDOP Along the Flight, Array 3

If neither flight profile nor receiver array are specified in advance, a desired region for a field of good GDOP can be defined that can be used as a criterion for array design. For example, such a field of good GDOP could be a fifty mile wide by fifty mile long by three mile high rectangular box of airspace starting at 5000 feet above the ground, with the GDOP inside the box no greater than six. This 'good GDOP region' concept also simplifies the choice of flight profile, because any flight profile that stays inside the good GDOP region is guaranteed to be acceptable, making the system more flexible and also more robust than it would be if tailored exclusively for a single flight profile. Hence, the global 'good GDOP region' concept has been extensively used in analyzing and comparing receiver arrays.

Comparing the GDOP fields of various arrays, a feasible array that is good for every intended application cannot be found. In free space, the inside of a spherical array does prove to have the best GDOP for a given number of receivers, as shown in Fig. 3.11. Even just four receivers in a tetrahedron provide GDOP values less than three throughout the entire region inside the tetrahedron. Such a region of low GDOP is hard to improve upon. Indeed, adding additional receivers to the sphere does little to improve the low GDOP region, as shown in Fig. 3.12. Unfortunately, it may not always be possible or desirable to configure the receivers in a spherical array. If it is necessary to navigate at some distance from the array, Figs. 3.13 and 3.14 show that a planar array might prove just as good as a spherical array, provided that there are at least ten to fifteen receivers in the planar array. In the case of the SARS, it is desired to make a ground based (almost planar) array have the same kind of low GDOP region that the satellite GPS provides, at least for a small volume of test range airspace.

This necessity for low GDOP at low altitudes forces severe design penalties on the SARS. For a ground based array, with an overall size somewhere around a fifty by fifty mile square, a full overflight by an aircraft may take it through variations of GDOP in the twenties to the hundreds, depending on the array and the flight profile. Such a large variation happens because the distances from the transmitter to the receivers are so small that the aircraft's motion drastically changes the angles on the line-of-sight (LOS) vectors drawn from the receivers to the transmitter. Obviously, this does not happen with satellite GPS. Since GDOP is a function of these angles, drastic changes in those angles can produce drastic changes in GDOP. To reduce the drastic changes in the GDOP, the receivers should be as far as possible from the transmitter, while still retaining enough angular variation from the transmitter's point of view to obtain good GDOP. Unfortunately this idea puts severe constraints on the design of a ground-based receiver array.

The behavior of the GDOP fields reveals severe constraints on the design of the ground based array. The two constraints are the horizontal extent of the array and the tradeoff between receiver density and minimum useful altitude of the system. The first constraint is the fact that for any ground based array of reasonable size for flight testing ($>$ 20 mi. diameter), the useable area of low GDOP will only occur directly above the array. Fig. 3.15 shows this constraint. Looking at the plot, it is evident that most of the low GDOP region is simply too high for aircraft to fly through. The part that is low enough to be useable is just directly over the array. So, in order to achieve a good fifty mile by fifty mile GDOP airspace, the

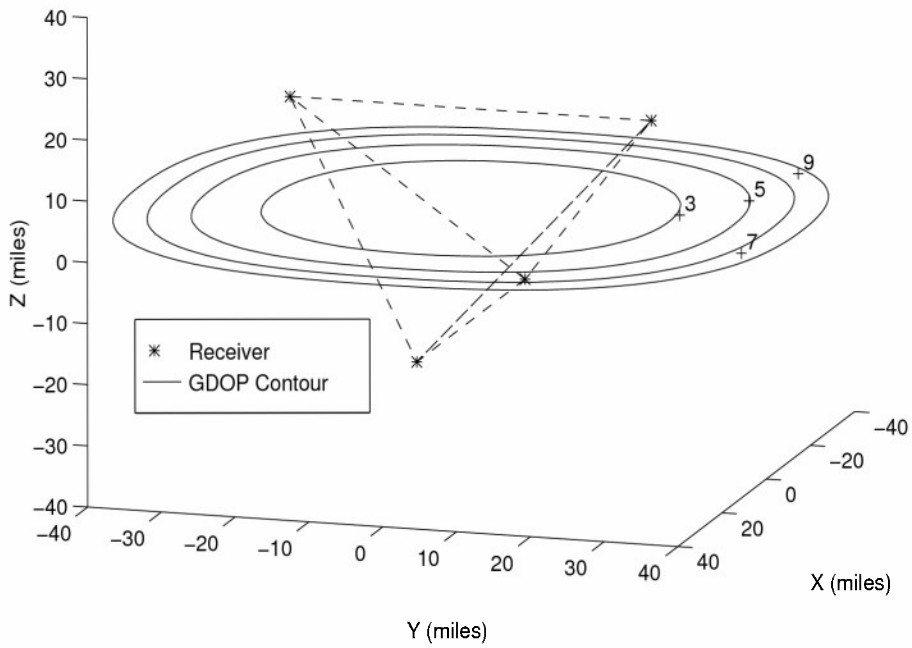


Figure 3.11 GDOP Fields for 4 Receivers in a Tetrahedron

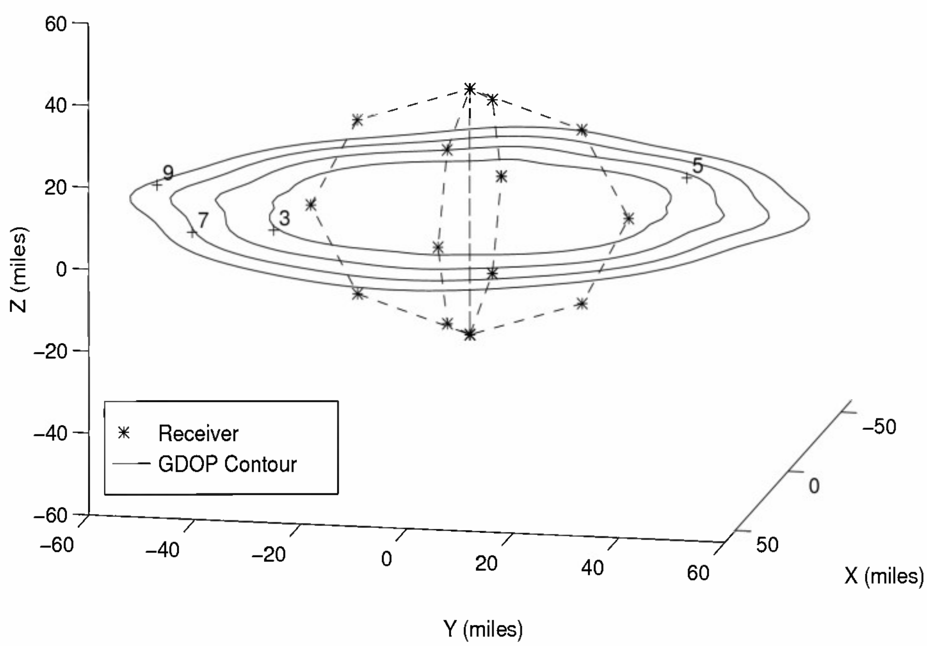


Figure 3.12 GDOP Fields for a 14 Receiver Sphere

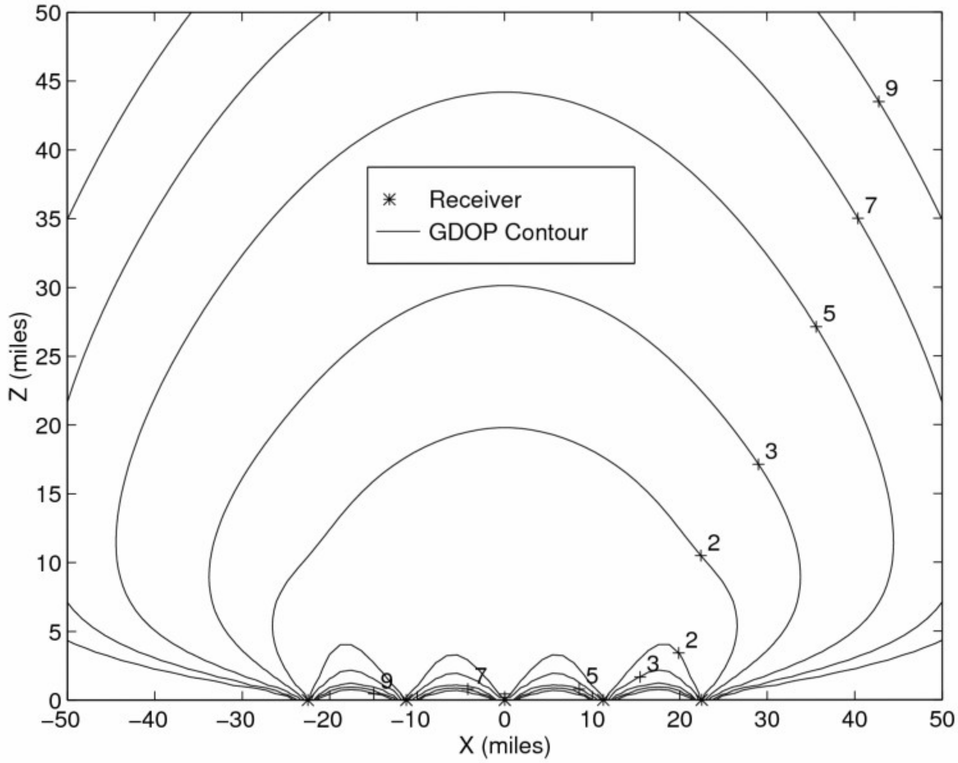


Figure 3.13 Vertical Cross Section of GDOP Fields, 25 Receiver Planar Grid

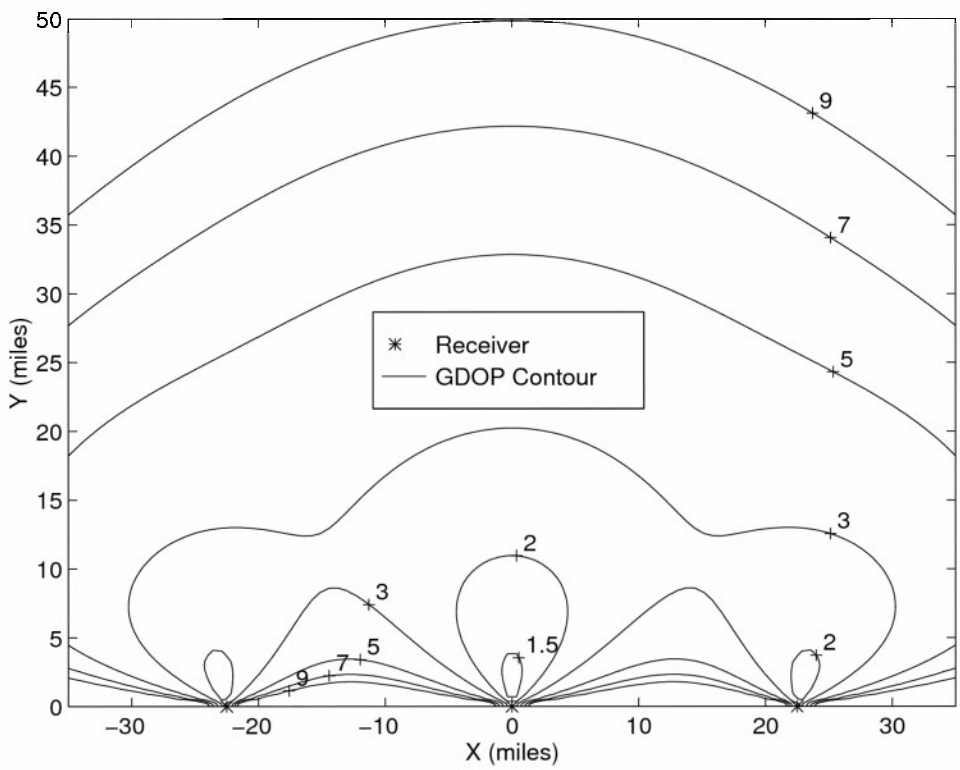


Figure 3.14 Vertical Cross Section of GDOP fields, 9 Receiver Planar Grid

receiver array must also occupy at least a fifty by fifty mile square. The second constraint is that there is a lower altitude limit for the low GDOP region. This limit is directly dependent on the density of the receivers on the ground, directly underneath the specified low GDOP region; *lowering this lower altitude limit means increasing the receiver density on the ground.* For example, an airspace with low GDOP (GDOP$<$ 5) that is thirty miles square at one mile above the ground level requires about twenty- five receivers. If the region needs to be larger, then the number of receivers must go up with the increase in array area. If the altitude lower bound needs to be lower, than the receiver density needs to be increased. Each way, the number of receivers needed increases much faster than the useful extent of the low GDOP region. A point of diminishing returns is rather quickly evident. The worst part of it is that above the minimum altitude, the benefit of the additional receiver density is almost negligible. Clearly, the higher the minimum altitude bound on the low GDOP region is, the fewer receivers are needed in the array, or the larger the array (and therefore the test airspace) can be made with the same number of receivers. Of course, if low GDOP is desired at the runway, then there will be serious problems. Even in mountainous country, and/or using 300 foot towers, it is hard to get the GDOP at any point within a few thousand feet of the array plane (ground) to be uniformly less than from fifty to one hundred.

The conclusion drawn is that achieving low GDOP at low altitudes over a wide area requires a large number of receivers.

Taken as a whole, the plotting of the fields of GDOP of receiver arrays clarifies the design issues for ground based GPS receiver arrays, like the SARS. First, the visualization of the GDOP airspace the transmitter would be flying through helps to show the flexibility available when both flight profile and receiver array can be specified. In fact, the concept of the field of 'good' (low) GDOP, which came about as a way to simplify the design process, proves useful as a general benchmark of array performance, removing the dependence on the flight profile, and adds flexibility

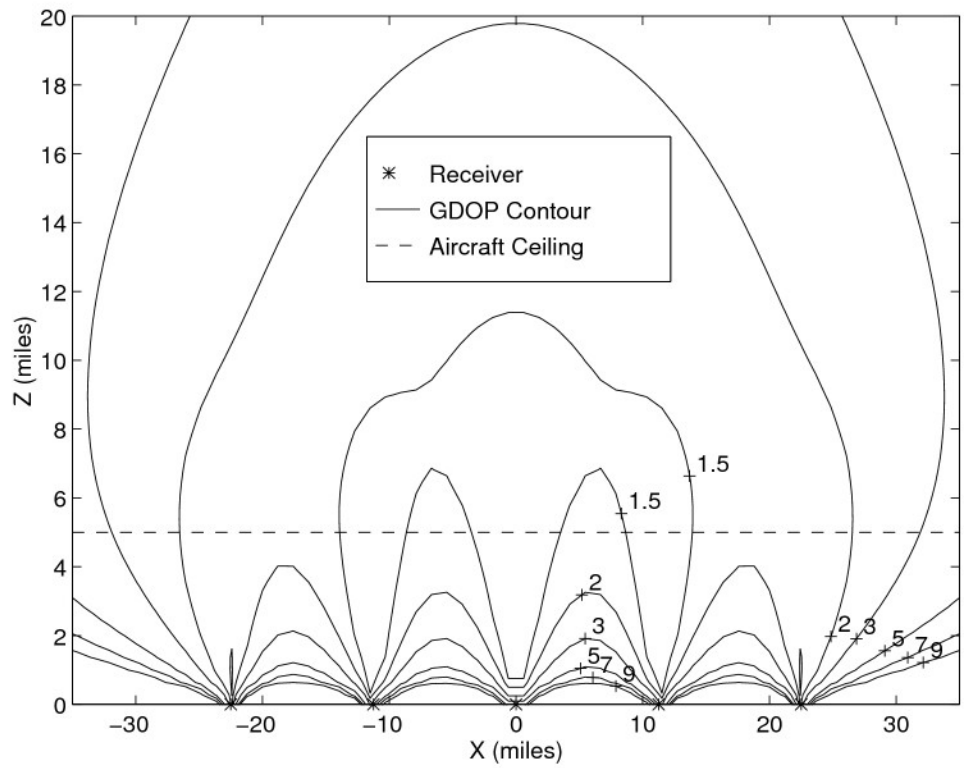


Figure 3.15 Vertical Cross Section of GDOP Fields, Useable Range

to the resultant system. Second, the comparison of the GDOP fields of different array configurations show that, while the best GDOP is found at the inside of a spherical array, some applications could be better off by using other configurations. Unfortunately, the SARS is not one of those applications, and low GDOP comes only at the cost of large numbers of receivers. Finally, the GDOP field plots of ground-based arrays show the tradeoffs between the extent of useable good GDOP airspace, the number of receivers, the receiver density, and the minimum altitude for low GDOP. These tradeoffs can now be used to optimize the ground-based receiver array.

3.1.3 Summary: Characterization of GDOP. As the first step towards the eventual goal of receiver array optimization, the behavior of GDOP around receiver arrays is characterized and important insights are gained. The analysis of the two dimensional navigation problem illustrates the tradeoffs inherent to the design of the SARS: receiver spacing vs. minimum transmitter altitude, and number of receivers vs. array size and GDOP. Although these results are not directly applicable to the SARS, they illustrate the basic design constraints inherent in optimizing the geometry of an inverted GPS receiver array. The plotting of the GDOP fields shows that the low GDOP region is a useful performance benchmark, that there is not one 'optimal' array configuration for all applications, and that ground based GPS receiver arrays like the SARS require many more receivers than the satellite GPS to achieve the same GDOP in a reasonable airspace. This is the result of the tradeoffs between minimum altitude, receiver density, and the size of the useable low GDOP airspace: the same tradeoffs found from the two dimensional analysis. Taken as a whole, the analysis of GDOP shows just how different the designs of the SARS and satellite GPS are. First, the SARS requires a large number of receivers to be useful. Second, the receivers must be placed carefully, with the flight profile in mind, or unacceptably high GDOP may result. Third, the SARS has a stringent minimum altitude below which the GDOP is terrible. These three differences make the design of the SARS totally different than that of the GPS satellite constellation, and make the choice of receiver array and flight profile much more difficult. The results of the characterization of GDOP clarify the array design issues, and provide insights that prove invaluable in the numerical optimization of the SARS' configuration.

3.2 Alternate Measures of Geometric Sensitivity

Although the GDOP is widely used in the literature as *the* measure of geometric sensitivity for GPS systems, it is desired to simplify the array design process by finding a better measure of GPS position sensitivity to measurement errors than the GDOP. Two candidates to be evaluated vs. GDOP are the condition number of the H matrix and and the angular field of view the receiver array takes up from the transmitter's point of view. The condition number of H is a candidate because it is a direct expression of how close the H matrix is to singularity. It is thought that cond(H) will prove to be a better measure of geometric sensitivity than the GDOP. The angular field of view (FOV) is a candidate because it it allows physical insight into the problem which may aid the array optimization. Although neither function is actually easier to compute than the GDOP, both functions bring more geometric insight into the problem.

3.2.1 The Condition Number of H. Since the geometry goes 'bad' when the H matrix approaches singularity, and a natural measure of how close a matrix is to singularity is the condition number, an investigation is conducted to determine whether the condition number of the H matrix provides a better measure of the geometric sensitivity than the GDOP.

The condition number of a matrix H is defined as the ratio of the largest and smallest singular values of H, viz.,

$$Cond(H) = \sqrt{\frac{\max \lambda(H'H)}{\min \lambda(H'H)}}$$

where $\lambda(\cdot)$ denotes the eigenvalues of the matrix. A high condition number indicates that the matrix is approaching singularity. A condition number of unity indicates that the matrix is as well conditioned as possible. Indeed, only an orthogonal matrix with equal eigenvalues (e.g., an orthonormal matrix multiplied by some scalar) can have a condition number of unity.

As mentioned before, an ill conditioned H matrix can cause the relative error in the computed position to be much larger than the relative measurement error. The upper limit on how much the position error could possibly be increased is given by the condition number, as shown by the following Theorem.

Theorem III.1 *Consider the matrix equation*

$$z = Hx$$

with x being the desired output and z being the input. Assuming the H matrix is error free, the relative error magnitude of the computed position is bounded above as

$$\frac{\| \Delta x \|}{\| x \|} \leq Cond(H) \frac{\| \Delta z \|}{\| z \|}. \tag{3.1}$$

Thus, the condition number of a matrix represents the worst case error amplification that could result from using the matrix to solve a linear system of equations. For the matrix A in the contrived example, the condition number was 10^6: the error amplification achieved its theoretical maximum. For the H matrix, since it embodies the GPS system measurement geometry, cond(H) shows the maximum possible impact of the measurement errors on the system accuracy due to geometry alone. The bounding property of the condition number on the error amplification makes cond(H) a robust measure of geometric sensitivity, more conservative than the classical definition for the GDOP function. This could prove useful for optimizing the receiver array geometry.

But just how much different is the information content of cond(H) from that provided by the classical definition of GDOP? Both functions provide some measure of geometric sensitivity based on the H matrix. If the classical GDOP function provides the same information as the condition number, then continued use of the classical GDOP figure of merit for the optimization is preferable. And if the two functions are different, it is desirable to know which one is better for the purposes of optimizing the receiver array geometry.

To see how cond(H) differs in information content from the classical GDOP, a Monte Carlo simulation experiment was conducted. The simulation first formed a receiver array, then created a 'blanket' of transmitter (user) positions randomly throughout a region enclosing the receiver array. The H matrix at each transmitter point was then generated, and the GDOP and cond(H) were taken. Running the Monte Carlo simulation a few times, it was found that, for a fixed receiver array, the classical definition of GDOP is roughly linearly correlated with the cond(H) function, as is shown for a 21 receiver planar disk array in Fig. 3.16. Thus, $a_1\text{cond}(H) \leq GDOP \leq a_2\text{cond}(H)$ for a given receiver array, where a_1 and a_2 are the slopes of the lower and upper bounds of the cone shaped point distribution on the correlation plots. As surmised, both functions convey somewhat similar information. But, as the simulation was run over different arrays some differences between the functions emerged.

Looking at the GDOP vs. cond(H) correlation plots for different array configurations and arrays of different numbers of receivers, it was found that the lower bound slopes a_1 of the correlation plots depend on the number of receivers in the array, and that the spread of the correlation plots $(a_2 - a_1)$ depends on the array's geometrical configuration. Fig. 3.18 shows the dependence of the correlation lower bound slope a_1 on the number of receivers. The spread of the correlation plot refers to the width of the cone shaped correlation point distribution; i.e., how big the difference is between a_1 and a_2. This spread is tighter, the better the array's geometry is. Fig. 3.16 and Fig. 3.17 show how the correlation plot spread changes with array geometry. The correlation plot of the 21 receiver spherical array (Fig. 3.17) shows a much tighter distribution of correlation points than that of the 21 receiver disk (Fig. 3.16).

From Fig. 3.18, one can see that the slope of the correlation lower bound decreases with an increase in the number of receivers in the array. This happens because the GDOP always decreases with the increase in the number of receivers, but the condition number may either decrease or increase, depending on whether the geometry is improved or not.

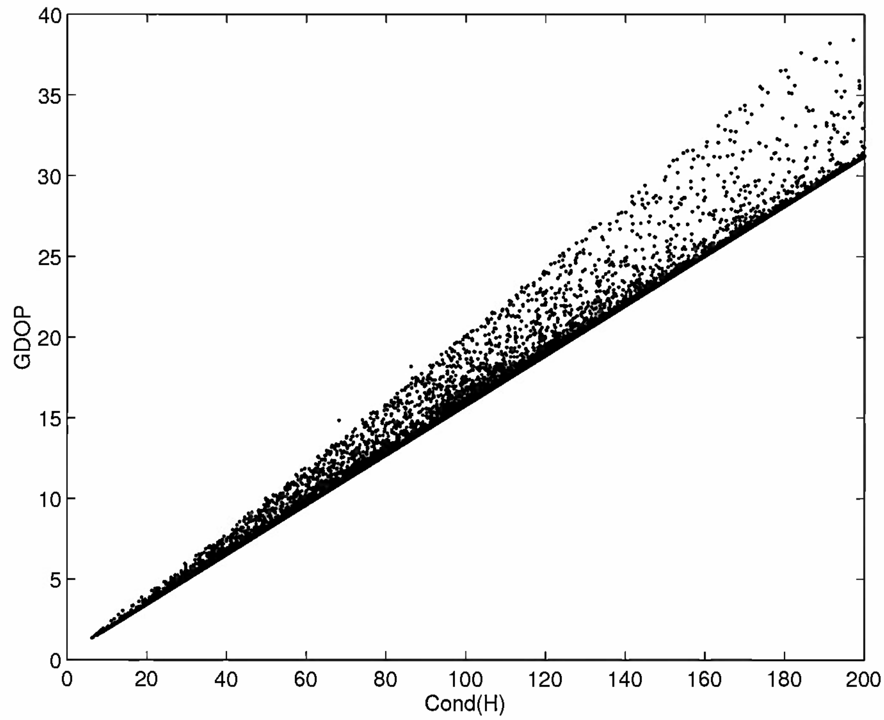


Figure 3.16 Correlation between GDOP and Cond(H), 21 Receiver Disk

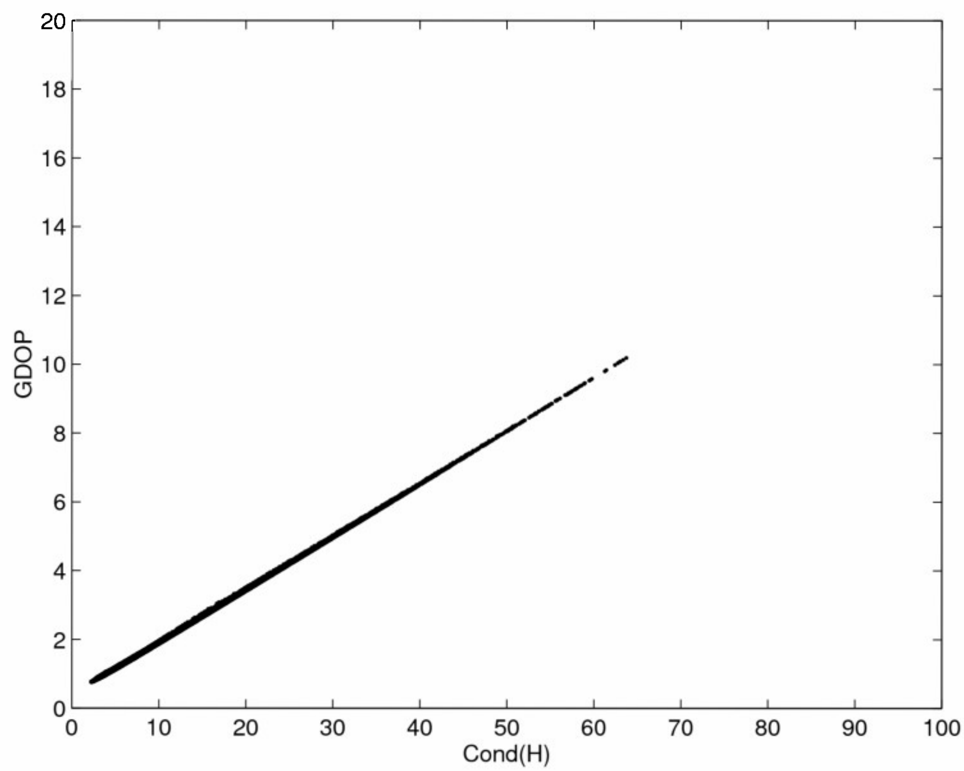


Figure 3.17 Correlation between GDOP and Cond(H), 21 Receiver Sphere

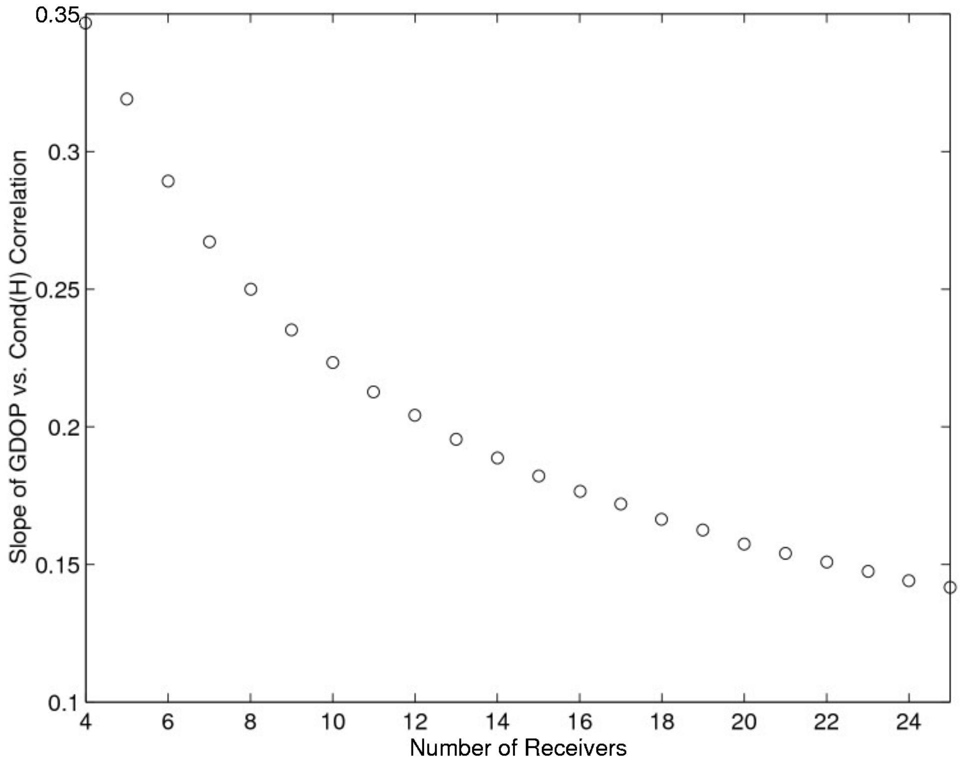


Figure 3.18 Slope of GDOP/Cond(H) Correlation Lower Bound vs. Number of Receivers

Hence, the GDOP parameter is determined by both the number of receivers and the array geometry, but the condition number is a measure of geometry alone; *this is the main difference.* Also, from the plots showing the differences in correlation spread, one can see that the GDOP and cond(H) functions are less strongly correlated in planar arrays than for three dimensional arrays with appreciable volume. In planar arrays, the GDOP at points with the same condition number can vary significantly. This effect occurs mainly near the plane of the array. Since the SARS' usable airspace will lie almost entirely within the region close to the plane of its array (the ground), the fact that cond(H) has a weaker correlation to the GDOP in this region may be important.

In conclusion, these results show that the condition number of the H matrix is a worthwhile candidate for use in place of the classical GDOP in the optimization of ground based (nearly planar) receiver arrays (e.g., the SARS) for three reasons.

1. The absence of direct impact of the number of receivers on the condition number of H allows the focus of the optimization to be on array geometry alone. Arrays with differing numbers of receivers should be able to be compared in terms of measurement geometry over the flight profile, without having to 'compensate' for the difference in numbers of receivers, as would be the case if GDOP is used. Of course, the more measurement taken, the better the SARS' accuracy will be, but that can be dealt with as a separate issue. Using cond(H) decouples the problems of finding good array geometry and deciding on the number of receivers for filtering purposes.

2. The condition number establishes an upper bound on the error amplification due to measurement geometry. Therefore, it can be used as a criterion for the largest allowable error magnification in the computed position, due to geometry. A way to use this concept is to design the receiver array to produce a region of airspace around a flight profile in which the cond(H) is less than a specified maximum bound. Given the relative independence of cond(H) on the number of receivers, the array could also be optimized with respect to numbers of receivers as well.

3. The behavior of the condition number of H at low altitudes over a planar array is somewhat different from that of GDOP. It has been found that using GDOP as the cost function for planar array optimization causes numerics problems. Since cond(H) behaves somewhat differently than the GDOP next to the plane of a planar array, it may be that cond(H) will make a more effective cost function for receiver array optimization.

These three reasons suggest that the condition number of H may prove more worthwhile for use as an optimization criterion than the classical GDOP. However, as the correlation plot shows, the projected improvements from using cond(H) instead of the GDOP are likely to be small, due to the high correlation between them. In fact, the main consideration between the functions may well be how well they 'lend' themselves to the array optimization process. Therefore, both functions are used as criteria to aid the receiver array optimization process. Which one is better is determined by how the resulting 'optimized' arrays turn out.

3.2.2 Field of View. The other alternate measure of geometric sensitivity investigated is the transmitter's angular field of view of the receiver array. The importance of this concept was noticed in the course of analysis on a simplified, two-dimensional (planar) GPS system. In that analysis, it was noticed that there was a strong correlation between the total angular field of view the receiver array took up from the user's vantage point and the GDOP. In fact, it was found that there was a limiting curve on the plot of the GDOP as a function of the angle of view. For every configuration of the array that produced an equivalent angle of view, the GDOP could never be lower than this bounding curve. This kind of geometric information could be used to 'eyeball' the size of the array necessary to produce a region of satisfactorily low GDOP, or used in place of the GDOP as a more easily visualized criterion for array optimization. Therefore, the three dimensional equivalent of the angle of view needed to be examined, to see if it could be as useful as in the two dimensional case.

To check out this field of view idea, it is necessary to calculate field of view (FOV) from the transmitter's and receivers' positions. Although readily visualized and easily calculated in two dimensions, the calculation of the FOV angle in three dimensions is not so readily done. First of all, the meaning of a three dimensional FOV needs to be clearly defined. Then, a means to compute the field of view needs to be found. Finally, once that's done, the analysis is carried out to determine if the 3-D FOV is a meaningful design tool.

The three dimensional FOV of the receiver array can be visualized as the portion of the transmitter's full field of view occupied by the convex region enclosed by the receiver locations. The transmitter's full field of view is defined as the surface area of the unit sphere: i.e., 4π. The field of view taken up by the receiver array is the surface area of the projection of the receiver array's convex hull onto the unit sphere centered on the transmitter. The projection is along LOS directions. A way to visualize this is as follows. Imagine the transmitter to be a point source of light, and the receiver array convex hull to be opaque. Imagine a sphere centered on the transmitter that is large enough to encompass the receiver array in its relative position with respect to the transmitter. Now, measure the surface area of the receiver array's shadow as projected onto the enclosing sphere. Divide the measured surface area by the total surface area of this enclosing sphere to get the fraction of the total FOV taken up by the receiver array. Finally, multiply this fraction by 4π to get the amount of the transmitter's FOV that is occupied by the receiver array.

For example: if the transmitter is inside the convex hull of the receiver array, then the array occupies the transmitter's entire FOV, which is 4π (the surface area of the whole unit sphere). If the transmitter is located on the surface of the receiver array's convex hull, the array would occupy half of the transmitter's total FOV, i.e., 2π. If the receiver is a significant distance outside the receiver array convex hull, then the FOV is only a fraction of 4π. (Inside a cave, the earth takes up a person's entire FOV. On the earth's surface, the earth takes up about half of a person's entire field of view. In space, the earth takes up less than half of the astronaut's total FOV.)

Now that the FOV has been defined, an algorithm to compute it needs to be designed. The following algorithm is used:

1. **Initialize the problem.** Change coordinates to place the transmitter at the origin. Move the receivers along line-of-sight directions until they lie on the unit sphere.

2. **Check to see if the transmitter is inside the receiver array's convex hull or not.** This check uses a control algorithm to search for a 'hemisphere axis vector.' If it exists, the dot products between this hemisphere axis vector and the positions of all the receivers in the array will be greater than or equal to zero. If such a vector is found, then all the receivers are in the hemisphere defined by the hemisphere axis vector and the plane passing through the origin that is perpendicular to this vector. If a hemisphere axis vector is not found, then the transmitter is assumed to be inside the receiver array's convex hull and therefore the FOV is 4π and the algorithm is done. If a hemisphere axis vector is found, then the vector is output to the next step in the algorithm.

3. **Sort the receiver array to identify those receivers on the projection of the convex hull** First, the hemisphere vector is used to rotate the coordinates so that all the receivers are in the $z \geq 0$ hemisphere. Second, the list of receivers is sorted to find the receivers on the perimeter of the projection onto the unit sphere. Those receivers on the perimeter are used to calculate the FOV.

4. **Calculate the FOV, given that the transmitter is outside the receiver array.** The area of the spherical polygon is calculated using the list of receivers found in the previous step. This is done by first breaking the polygon into triangles, then using the spherical Heron formula to calculate the area of the spherical triangles. The areas of the spherical triangles is added up to produce the amount of the transmitter's FOV taken up by the receiver array.

Although the algorithm is slow, and uses an iterative technique for determining if the transmitter is inside the convex receiver array volume, it suffices for the purposes of this research.

Finally, now that a means of calculating the field of view is on hand, the correlation between GDOP in three dimensions and the 3-D FOV can be examined. As with the GDOP vs. cond(H) correlation study, a Monte Carlo simulation in Matlab is created to perform this task. The simulation works as follows: First, a receiver array is generated. Then a large number of candidate transmitter locations (10000+) are randomly generated over a region enclosing all the desired locations for the transmitter. This region usually was made to enclose the receiver array as well as all the airspace outside of the receiver array in which the GDOP was good. Next, the GDOP and FOV are evaluated at each one of these transmitter locations. Finally, the GDOP vs. FOV at all these points are plotted, making 'scatter diagrams.' The scatter diagrams are correlation plot of GDOP vs. FOV for specific receiver arrays. They are evaluated to see if the FOV has enough information content to be a useful measure of array geometry for array design purposes.

The analysis shows two things: the 3-D GDOP vs. FOV correlation lower bound is similar to the 2-D case and the patterns on the correlation plots vary with different array configurations. Unfortunately, a critical look at both results show that the FOV concept is not very useful as a design tool for ground-based receiver arrays, for the following reasons.

The first result of the FOV study, as shown in Fig. 3.19, is that the lower bound curve of GDOP vs. FOV holds true in three dimensions. The FOV for a three dimensional array does give a lower bound on the possible GDOP, regardless of array configuration. The simulation does not produce a closed form function for the lower bound, but a lookup table can be generated from the simulation data. Unfortunately, though, this lower bound on GDOP is not very useful for the SARS, due to the severe 'choppiness' of the GDOP as the transmitter approaches the plane of the receiver array (as is the case during most of a typical flight profile). The GDOP peaks producing the choppiness in the regions of high FOV are well illustrated by Figs. 3.19 and 3.21. Indeed, if the transmitter is directly above a receiver, the GDOP

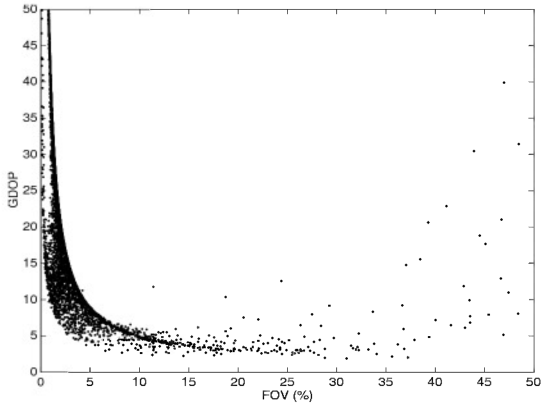


Figure 3.19 GDOP vs FOV, Array 1

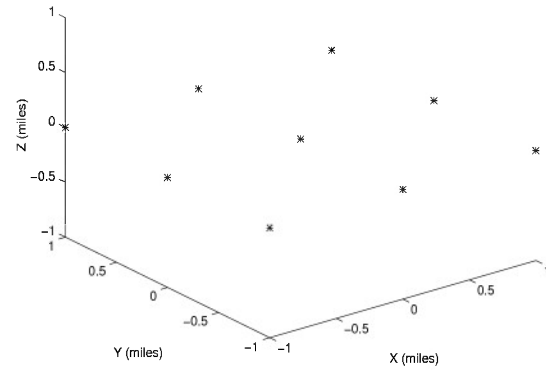


Figure 3.20 Array 1

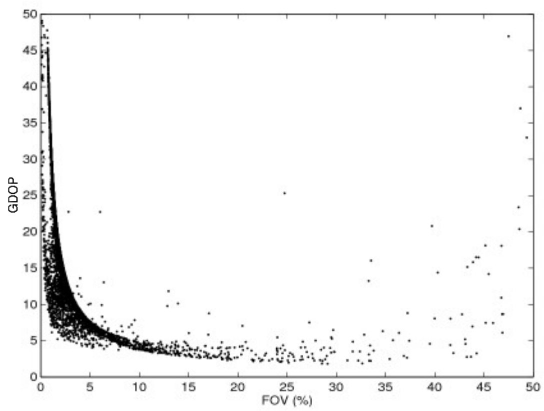


Figure 3.21 GDOP vs FOV, Array 2

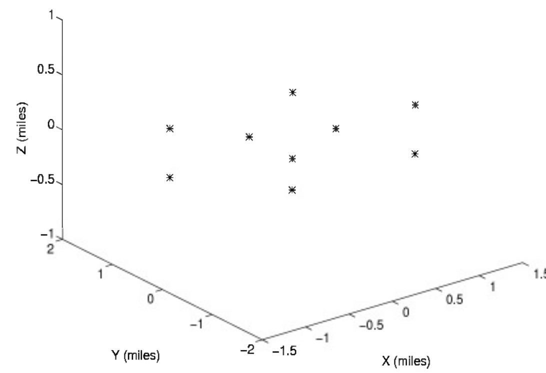


Figure 3.22 Array 2

is very low (around two), but if the transmitter is not directly above a receiver, the GDOP at the same altitude could run very high (around ten to twenty). Close to the surface of a planar array, it's the maximum GDOP that's critical, not the minimum, and, unfortunately, the FOV gives no information on the maximum GDOP. Hence the FOV concept is just not very useful for designing a planar array with good GDOP at low altitudes.

The second result of the FOV study is that the GDOP vs. FOV correlation pattern is different for different receiver array geometries. Grid arrays have different correlation plots than random arrays. Planar arrays show different correlation plots than spherical arrays, etc. Looking at the correlation plots of Figs. 3.19-3.22, it's not clear which configuration is better. The best GDOP comes at different FOV, and the degree of GDOP choppiness varies as well. There is information in the GDOP vs. FOV correlation plot, but it is in a form that is difficult to use. The parametric nature of the GDOP vs. FOV plot disguises what kind of airspace of good GDOP the arrays might be producing. So the differences in the GDOP vs. FOV correlation plots have little usefulness for the SARS array design.

3.2.3 Summary: Alternate Measures of Geometric Sensitivity. As stated above, the results of the examination of alternate measures of geometric sensitivity show that the condition number of the H matrix proves to be a useful concept in receiver array design, but the field of view concept does not. Cond(H) provides the same general information as the GDOP, but has differences which make it especially useful as a measure of geometry for receiver array optimization. The first difference is that cond(H) measures the effect of geometry alone, decoupling the problem of finding good array geometry from that of choosing the number of receivers needed for filtering purposes. The second difference is that cond(H) provides a worst case limit on error magnitude amplification due to measurement geometry, a useful concept for array design. The third difference is that cond(H) may prove more amenable to numerical optimization than the classical GDOP. These differences make cond(H) a viable candidate for use in the receiver array optimization cost functions. The field of view, on the other hand, is not good enough to be useful for the design of ground based GPS systems, due to the problem of choppy GDOP and the difficulty of comparing the GDOP vs. FOV plots for different arrays. Although the FOV concept does provide qualitative insight into GPS array design, it is just not reliable enough to be of much use. In conclusion, both the GDOP and cond(H) will be used as measures of the geometric sensitivity in the optimization of the SARS' receiver array, but the FOV will not be used. The results of the optimization will determine which measure, GDOP or cond(H), is better for optimizing the SARS.

3.3 Geometric Sensitivity Cost Function

Now that the three alternate measures of geometric sensitivity (GDOP, cond(H) and FOV) have been discussed, the choice of which one to use in the cost functions in the receiver array optimization programs needs to be made. The previous analysis indicates that there are two possibilities: GDOP and cond(H). Many trial cases were tested with both GDOP and cond(H) during the development of the array optimization routines. The results are interesting. As the correlation plots between GDOP and cond(H) from the cond(H) analysis of Section 3.2.1 indicate, there is not much difference between the GDOP and cond(H) for a fixed number of receivers. The optimized arrays with respect to both turn out rather similar in shape, as is expected. However, the optimization runs using GDOP are much more sensitive to initial receiver array configurations than those runs where the cond(H) is used instead. The GDOP seems to make a cost function that is less conducive to numerical optimization than the cond(H). The optimization runs using GDOP are very likely to 'prematurely converge,' or get stuck on some small minimum in the cost function that is nowhere near the global minimum. What this means is that using the cond(H) in the cost function of the array optimization produces arrays that have better cond(H) *and* GDOP than if the GDOP were used instead.

This result is very surprising. Numerics difficulties with the GDOP as a cost function make it better to use the condition number of H for at least the initial optimization of a receiver array with respect to both cond(H) and GDOP. Of course, since the GDOP and cond(H) are so highly correlated, an array that produces generally good cond(H) will produce good GDOP, and vice-versa. The difficulty is in finding the good receiver array to begin with, and using cond(H) in the cost function has a better chance at finding it.

The cond(H) is attractive for reasons other than just numerics. Section 3.2.1 lists two additional reasons for using cond(H) First, using cond(H) decouples the problem of array geometry determination from that of determining adequate number of measurements for filtering purposes. The desired number of receivers for filtering purposes can be chosen, then a configuration for them can be determined using cond(H) that minimizes geometric sensitivity. Second, the condition number of H can be used directly in the design process to specify a worst case accuracy of the system.

Therefore, the condition number of H is used as the measure of geometric sensitivity for the receiver array optimization problem.

3.4 The Receiver Array Geometry Optimization Problem

The receiver array geometry optimization problem is to find the locations of the receivers and transmitter that will minimize the geometric sensitivity of the GPS position calculations for the SARS. There are several things that need to be done to properly set up the problem: choose the independent variables, choose the cost function, and identify the relevant constraints. Once the problem is properly set up, a simple four receiver analytical thought experiment is conducted in order to get a feel for what kind of results should be expected from the optimization. The problem formulation is by far the most important step of the optimization process.

3.4.1 Independent Variables. There are two possibilities in choosing the independent variables for the optimization. The independent variables can be either the locations of some or all of the receivers, or the desired transmitter position or trajectory. One choice is for the transmitter profile points to be the independent variables. In the geometric sensitivity analysis of Section 3.1.2, it is found that a good flight profile could easily be found for the transmitter with respect to a fixed receiver array by simply plotting the cost function (yet to be chosen) in three dimensions with respect to transmitter location. The transmitter's flight profile would be determined by 'flying' the aircraft (the transmitter) through regions of low cost. Though simple, this method has one big disadvantage. It requires a good receiver array to be known, but does not help in finding it. The other choice is to make the receiver locations the independent variables. This creates a much more difficult problem, because there is no nice, simple way to visualize and find the best locations to put the receivers. One simply can't plot (or visualize) a cost function over the twelve or more dimensions which represent the available degrees of freedom associated with the receivers' locations. However, numerical optimization routines exist which can handle the high dimensionality of this problem, and can be used to solve it. Therefore, the independent variables for the numerical optimization were chosen to be the receiver locations.

3.4.2 Cost Function. The next step in the optimization problem formulation is the choice of cost function. Results from the geometric sensitivity analysis indicate that the best indicator of geometric sensitivity is the condition number of the visibility matrix H, rather than the standard Geometric Dilution of Precision (GDOP) metric used today in GPS work as a figure of merit. The H matrix is composed of the normalized projections of each pseudorange vector along each of the coordinate axes, (x, y, z, t), where t represents the time dimension. This is a (non orthogonal) transformation of coordinates from a basis of pseudorange vectors to a four dimensional Cartesian coordinate basis. This DCM transformation can be expressed in terms of component angle cosines, as shown in the following equation (for four receivers):

$$H = \begin{bmatrix} \cos\theta_{1x} & \cos\theta_{1y} & \cos\theta_{1z} & \cos\theta_{1t} \\ \cos\theta_{2x} & \cos\theta_{2y} & \cos\theta_{2z} & \cos\theta_{2t} \\ \cos\theta_{3x} & \cos\theta_{3y} & \cos\theta_{3z} & \cos\theta_{3t} \\ \cos\theta_{4x} & \cos\theta_{4y} & \cos\theta_{4z} & \cos\theta_{4t} \end{bmatrix} \tag{3.2}$$

Now, the condition number of the H matrix is the ratio of the largest singular value of H to the smallest singular value of H. Since the H matrix is composed entirely of component angle cosines, the condition number is a function of the angular orientation of the pseudorange vectors, not their magnitudes. This means that the actual units used in the array optimization are arbitrary, and therefore a working design could be scaled up or down if necessary. It also means that the transmitter position or trajectory is embedded in the cost function, as the pseudorange angles depend on the transmitter location as well as the receiver array configuration.

The appearance of the transmitter position in the cost function raises the question of how to account for movement of the transmitter along a flight profile. The method used in this research is to discretize the transmitter's trajectory and to evaluate the condition number at the discrete points along it. Then the mean, or the maximum, condition number is taken to be the cost function. Two questions are raised by this method. The first question is which operation on the list of discrete condition number points, averaging or taking the maximum, should be used to form the cost function. The second question is how might the transmitter flight profile influence the outcome of the optimization.

The choice of mean or maximum sets the tone for the optimization. The mean produces a cost function that is better behaved but obviously not as conservative as that produced by taking the maximum. Taking the mean is good for obtaining a smooth, (relatively) well behaved cost function, but may allow 'spikes' of high cond(H) along the flight profile. Taking the max avoids the cond(H) spikes (sometimes), but is not numerically as well behaved, so the optimization runs the risk of 'hanging up', converging to a local minimum in the cost function which produces an array that is obviously neither 'good' nor an expected outcome of the optimization. Experience shows that it is better to use the mean(cond(H)) in the gradient based optimization routine because of this risk. Also, if 'spikes' of high cond(H) are not a problem, the array optimized with the mean(cond(H)) will likely have larger regions of low cond(H) than the array optimized with the max(cond(H)). Of course, if the max(cond(H)) is truly what is desired to be minimized, then the max(cond(H)) should be used in the cost function. In this case, the optimization routine may require more carefully designed initial receiver locations and flight profiles in order to work properly.

The choice of flight profile also tends to impact the effectiveness of the optimization. Certain profiles may cause difficulties with certain optimization methods (gradient based or Monte Carlo). The two most glaring examples are profiles that cross their paths, e.g., figure eight shaped, and profiles that attempt to fill entire regions in order to produce low cond(H) inside the entire region, rather than just along a specific flight profile.

The profiles with path crossings tend to foul up the gradient based optimization routine. The gradient based optimization routine has trouble making the receivers cross the 'shadow' (projection onto the ground) of the flight profile. This can lead to trouble if the profile's shadow has several crossing points, because the profile may 'wall off' receivers from moving to the best locations during the optimization process. This effect is more pronounced when the discretization is fine. A (very) coarsely discretized profile may allow the receivers to slip between the points' projections onto the ground and affect the required crossings, thus reducing the risk of 'hangup'. Unfortunately, having the profile very coarsely discretized tends to produce spikes of cond(H) in between the points. The best approach seems to keep the fine trajectory discretization and do multiple runs with different initial conditions in an attempt to allow the receivers to go to the best places. Strictly speaking, that is a method that overcomes the trapping of the solution into a local optimum.

The profiles that attempt to fill whole regions trouble both the gradient and Monte Carlo optimization routines. Such regions can be built into the cost function by making the flight profile into a grid of points of constant altitude, with the altitude being the minimum altitude desired for the low cond(H) region. This type of profile runs the risk of asking too much from a given number of receivers. If there aren't enough receivers to cover the ground under the desired region, no placement can cure the problem, and the optimizer most likely will not do any good. The resulting arrays are usually spread too thin and are unsatisfactory. If this happens, it must be realize that a trade off is needed, and either the size of the desired region must be reduced or more receivers must be used in the array. However, even if there are enough receivers to potentially cover the desired region, this kind of profile produces a horrendous cost function that tends to 'hang up' gradient based searches, if the wrong initial conditions are used. This kind of profile is difficult to use with the optimization routines. If it is desired, though, several runs using both gradient and Monte Carlo techniques should be used, to provide data from which to create a good array.

3.4.3 Constraints. There are many constraints imposed by the 'real world' on the receiver array optimization problem. The constraints are receiver altitudes (imposed by the terrain), receiver motion, allowed receiver and transmitter locations, the boundary of the test range, and the number of receivers. In the fully constrained system, i.e., the SARS, the receivers must be located on the ground in fixed positions. The transmitter cannot travel below the ground or above its carrier aircraft's altitude ceiling, and the number of receivers has an upper limit. Although the fully constrained optimization is the desired end product of this effort, it has proven useful to consider the unconstrained optimization as well, to better understand the problem, and to see what effect the constraints have on the resulting receiver arrays.

3.4.4 Simple Thought Experiment. To shed light on the optimization problem, a simple experiment with four receivers and a fixed transmitter location is conducted.

Let the transmitter be located at the origin, and the four receivers constrained to lie on the unit sphere which encloses the transmitter. This can be done without loss of generality because the H matrix is a function of only the angular orientation of the pseudorange vectors. Let the four receivers' positions be denoted $R1$,$R2$,$R3$,$R4$, with $R_i = (X_i, Y_i, Z_i)$ being the vector from the transmitter (at the origin) to the receiver i. The angle component between the ith pseudorange vector and the jth coordinate axis is denoted θ_{ij}. Since the distance vectors from each receiver to the transmitter are all of unit magnitude, the angle cosines in the H matrix can be simplified, as shown in the following expression. In this case, the H matrix becomes a function of the positions of the receivers R_i on the unit sphere. The H matrix is formed as follows:

$$H = \begin{bmatrix} \cos\theta_{1x} & \cos\theta_{1y} & \cos\theta_{1z} & \cos\theta_{1t} \\ \cos\theta_{2x} & \cos\theta_{2y} & \cos\theta_{2z} & \cos\theta_{2t} \\ \cos\theta_{3x} & \cos\theta_{3y} & \cos\theta_{3z} & \cos\theta_{3t} \\ \cos\theta_{4x} & \cos\theta_{4y} & \cos\theta_{4z} & \cos\theta_{4t} \end{bmatrix} = \begin{bmatrix} X_1 & Y_1 & Z_1 & 1 \\ X_2 & Y_2 & Z_2 & 1 \\ X_3 & Y_3 & Z_3 & 1 \\ X_4 & Y_4 & Z_4 & 1 \end{bmatrix} \tag{3.3}$$

It is desired to find the receivers' positions R_i that minimize the condition number of H. For this simple case, it is possible to obtain the optimal solution. The key to the solution is to use the properties of the condition number to simplify the problem.

The condition number of an orthogonal matrix with column vectors of equal magnitudes is the lowest possible condition number, 1. Looking at the H matrix, one can see that it is impossible for all four columns to have equal magnitudes, but it may be possible for the columns to be orthogonal. Looking at the high degree of freedom, it seems likely that it can be done.

If the H matrix can be orthogonalized, the magnitudes squared of the columns will be the singular values of H, allowing the singular values of H to be expressed in terms of the columns of the H matrix. This allows the singular values of H (and hence the condition number) to be expressed as simple functions of the receiver positions, considerably simplifying the problem. This works because of the properties of orthogonal matrices. The singular values of a matrix H are defined as the square roots of the eigenvalues of $H^T H$. If the matrix H is orthogonal, $H^T H$ is a diagonal matrix, because all the non-diagonal terms (dot products between orthogonal columns) are zero. The eigenvalues of this diagonal matrix are the elements on the diagonal, which are the squares of the magnitudes of the columns of H. Therefore, if H can be orthogonalized, its singular values will be the squares of the magnitudes of its columns.

The condition number of a matrix can be expressed as the ratio between its maximum singular value and minimum singular value. Clearly, all the singular values of H should be as close as possible to being equal, to keep the ratio as small as possible. It would be best if all the singular values of H could be equal, but that is not possible. because the magnitude of the fourth column of H is always larger than the magnitudes of the other columns. The best that can be achieved is an orthogonal H matrix with the first three columns having equal magnitudes. This will produce the lowest condition number possible for the H matrix.

Let us suppose that everything works out: the columns are orthogonal, the receivers still lie on the unit sphere, and the magnitudes of the first three columns are equal. The magnitude of the first three columns can be solved for, and the minimum condition number possible for the H matrix can be found. The first step is to solve for the unknown magnitude of columns one through three, denoted here as C. This is done by summing the magnitudes squared of the first three columns, and noticing that the terms can be rearranged to yield the following equation.

$$3C^2 = \sum_{i=1}^{4} X_i^2 + \sum_{i=1}^{4} Y_i^2 + \sum_{i=1}^{4} Z_i^2 = \sum_{i=1}^{4}(X_i^2 + Y_i^2 + Z_i^2)$$

Since the receivers must lie on the unit sphere, the right hand side of the equation reduces to $\sum_{i=1}^{4}(X_i^2 + Y_i^2 + Z_i^2) = (1 + 1 + 1 + 1) = 4$. Therefore, $3C^2 = 4$, and the magnitude of each of the first three columns turns out to be $C = \sqrt{4/3}$. Since the columns are orthogonal, and the first three columns have equal magnitudes, the condition number is as low as it can get. The condition number for this postulated configuration is given by:

$$Cond(H) = \sqrt{\frac{\max \sigma(H'H)}{\min \sigma(H'H)}} = \sqrt{\frac{4}{4/3}} = \sqrt{3}$$

The above derivation is not yet a proof, but it does indicate that $\sqrt{3}$ may well be the best possible condition number of the H matrix, given four receivers. In fact, the number of receivers does not change the result, as long as there are at least four. This implies that the lowest possible condition number for the H matrix may well be $\sqrt{3}$, independent of the number of receivers in the array.

The backbone of this derivation is that an H matrix could be found that could satisfy all the conditions: receivers on the unit sphere, orthogonality, and the first three columns having equal magnitudes. Although the above constraints on the H matrix seem restrictive, it turns out that a matrix satisfying all the conditions is easily found, as shown below:

$$H = \begin{bmatrix} \frac{-1}{\sqrt{3}} & \frac{-1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & 1 \\ \frac{-1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{-1}{\sqrt{3}} & 1 \\ \frac{1}{\sqrt{3}} & \frac{-1}{\sqrt{3}} & \frac{-1}{\sqrt{3}} & 1 \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & 1 \end{bmatrix} \tag{3.4}$$

This matrix represents an equilateral tetrahedron of receivers centered on the transmitter, as shown in Fig. 3.23. Numerical optimization applied to this problem indicates that $\sqrt{3}$ is in fact the minimum cond(H) possible and that the equilateral tetrahedron is what produces this 'optimal' condition number.

This conceptual experiment has shown the basics of the optimization problem for a receiver array using the condition number of H as the cost function. Adding receivers, constraints, and multiple transmitter points to the optimization greatly in-

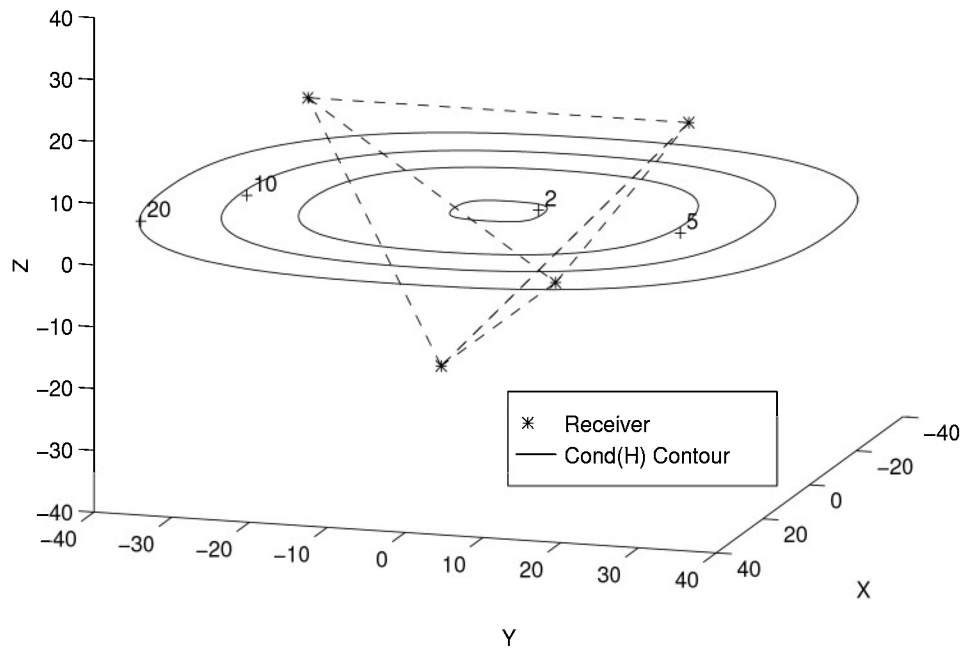


Figure 3.23 4 Receiver Array, Optimized about the Origin

creases the complexity of the problem, motivating the use of numerical optimization techniques to better attack the problem.

3.5 Receiver Array Optimization Tools

Several receiver array optimization tools are used to aid in the optimization of the SARS. These tools are: a graphical technique for evaluating receiver array geometry and two numerical optimization techniques used to optimized the receiver arrays. The numerical techniques are a Monte Carlo search program for good receiver locations, and an array optimization program using Sequential Quadratic Programming in the form of **constr.m** from the MATLAB Optimization Toolbox. These optimization tools are used to gain insight into, then solve the receiver array optimization problem.

3.5.1 Graphical Analysis of Array Geometry. The first optimization tool developed is for graphically evaluating receiver array geometry. The output of this technique has already been seen in Section 3.1.2. This tool takes a receiver array and a cost function, then generates a 'field' of cost function values as a function of transmitter location. Plotted around the receiver array, these fields make it easy to evaluate good or bad array geometry. An example of the output from this tool is Fig. 3.24.

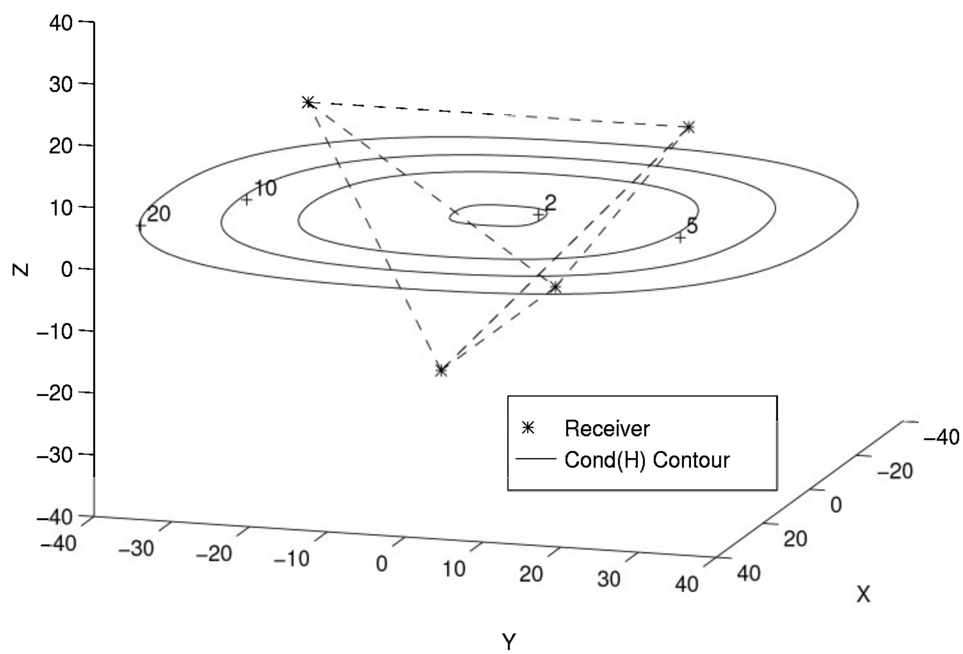


Figure 3.24 Cond(H) Fields for 4 Receivers in a Tetrahedron

This tool enables the array designer to evaluate the geometry of a receiver array at a glance. It also allows the designer to quickly hand-pick flight profiles for the transmitter that will lie within the regions of good measurement geometry. This graphical technique is a real help in receiver array optimization: it speeds up the task of array evaluation, gives insight into what makes 'good' array geometry, and is used heavily throughout the array optimization research to make sense of the results.

3.5.2 Numerical Optimization Techniques. Several numerical optimization techniques have been used in this research. The basic techniques used were a Monte Carlo search routine, a constrained optimization routine from the MATLAB Optimization Toolbox, and combinations of the two. All the numerical optimization routines worked to some degree, but rarely was the cost function well behaved enough to allow a globally optimal receiver configuration to be found. Therefore, these optimization tools were used mainly to develop insight into the problem of receiver placement and to develop 'rules of thumb' of good receiver array design, rather than to pursue an elusive global 'optimal' receiver placement.

Before the detailed discussion of these techniques, one question needs to be addressed: How does one evaluate the performance of a receiver array optimization technique? Over the course of this research it has been found that there are several criteria that can be used to evaluate the performance of the array optimization techniques. These criteria are:

1. Consistency: Are experiments repeatable? Do optimized arrays exhibit similar traits? Does the optimization consistently improve the arrays?
2. Symmetry: Is symmetry preserved/created in the optimized array?
3. Sensitivity to Initial Array Configuration: To what degree does the configuration of the optimized receiver array depend on the initial array put into the optimization routine?
4. Effectiveness: Are the optimized arrays satisfactory? How often do 'hang ups' occur? How likely is it that better arrays exist?

Using these criteria, the performance of the two numerical optimization techniques can be evaluated.

3.5.3 Monte Carlo Search. The Monte Carlo technique is essentially a random search for the best receiver array configuration. Random search is used to avoid the solution being 'hung up' on one of the (many) local minima of the cost function, which has been a significant difficulty in the optimization of planar arrays. The technique is as follows:

1. **The initial array is formed.** Numerical optimization requires a starting point for all the independent variables.
2. **The transmitter trajectory for the cost function is specified.** (A list of transmitter points)
3. **The desired number of iterations is set.** This parameter is the stopping criterion for the program. The number of iterations tells the program how many times to go through the whole receiver array, moving receivers around one by one to look for receiver locations that lower the cost function. The number of iterations influences how good the receiver array will be and determines how long the program will take. Usually, it takes two iterations (i.e., moving every receiver around twice) to converge to a reasonably good receiver array. More 'polished' results are obtained with six or seven iterations. As this program takes hours to run, it is better to set the number of iterations on the high side, to ensure complete convergence of the receiver array. (That way it can be started in the morning, forgotten about, and will be sure to have good results in the evening.)
4. **Generate alternate arrays.** One receiver is taken out of the array, and alternate positions for that receiver are randomly generated within a desired region, i.e., within specified bounds inside a region in space or in a plane. This produces many different alternate receiver arrays, differing only by the position of one receiver. It has been found useful to change the size of the region in which the selected receiver can move during the course of the optimization. The most general implementation would be to allow every receiver to move within the entire space over the course of the entire optimization, but that would be

very time consuming. Good results (and much shorter run times) are obtained by reducing the region of movement for reach receiver as the optimization progresses. This results in a coarse, medium, and fine optimization being performed on the receiver array. This method has obtained consistently good results. If the problem were to change, though, the region sizes for the three optimization stages need to be adjusted to reflect the new problem.

5. **The cost function is evaluated.** For each alternate array, the cost function is evaluated, and the array with the lowest associated cost becomes the new array.

6. **Repeat from step four.** Go back to step four and repeat, choosing the next (different) receiver from the array. Repeat until the iteration limit is reached. The receiver positions after all the iterations become the 'optimized' receiver array.

Research indicates that this random search technique does in fact improve upon the initial array with respect to a given flight profile. An example of this performance improvement is shown for a 25 receiver array in Figs. 3.25 and 3.26.

The performance achieved by this method is evaluated by considering the optimization performance criteria: consistency, symmetry, sensitivity to initial array configuration, and effectiveness.

Consistency:

The random search technique proves to be fairly consistent. Repeatability is not exact, but the optimized arrays for a given transmitter profile exhibit similar traits, and the cond(H) traces over the given flight profiles are consistently improved.

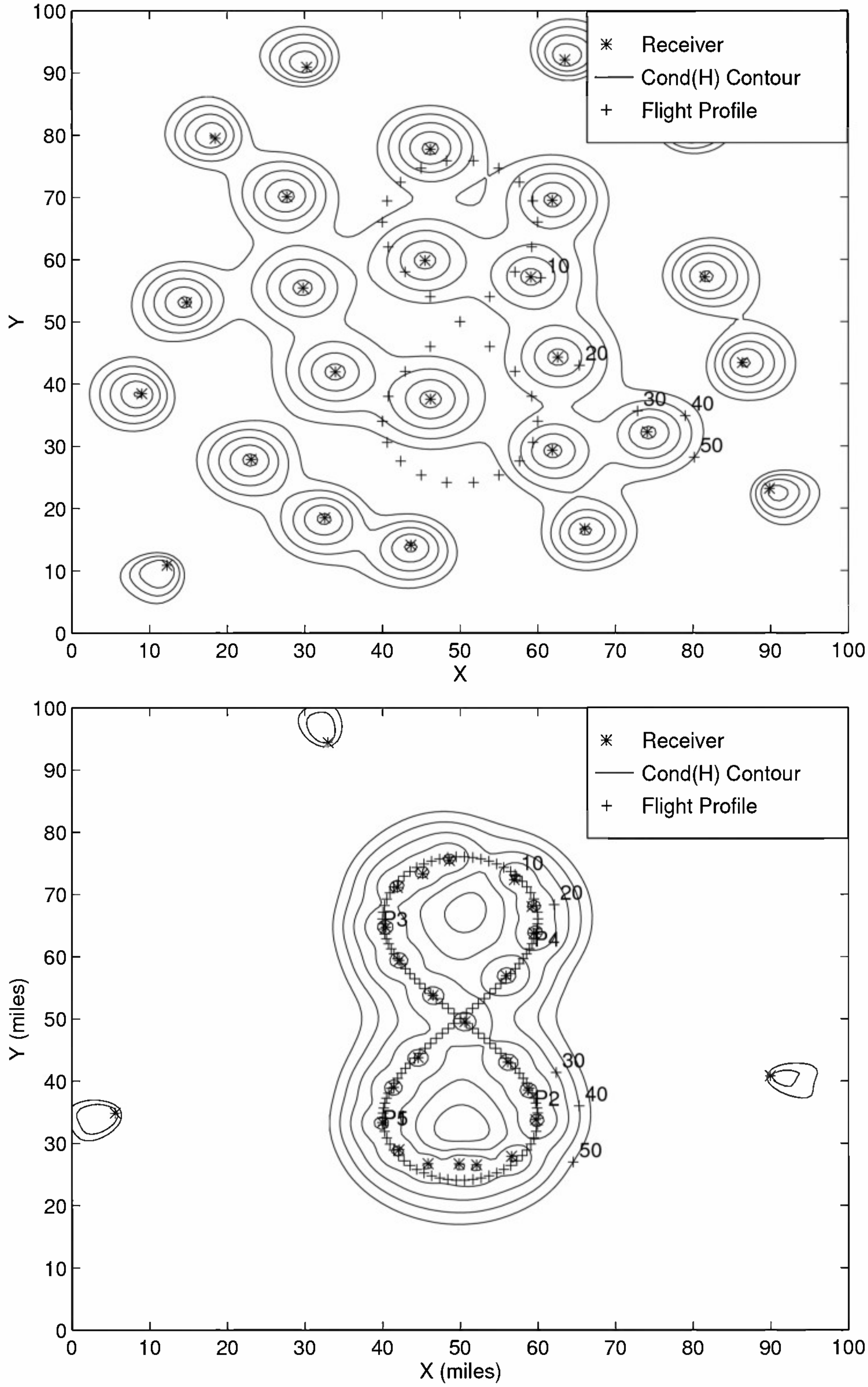


Figure 3.25 Initial and Final Receiver Arrays

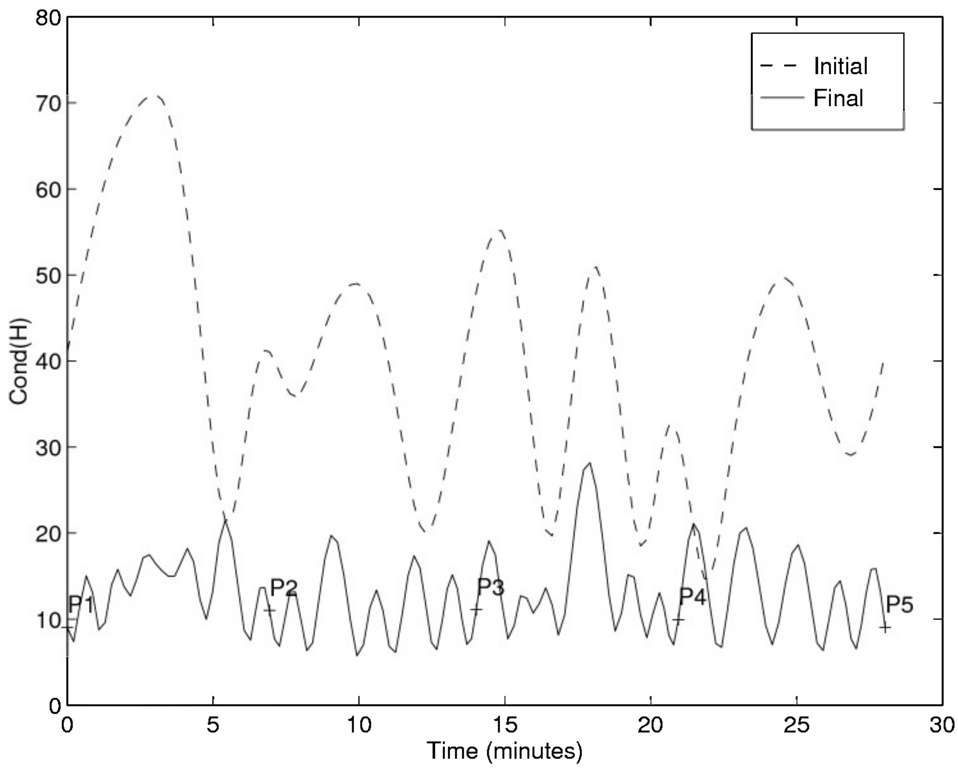


Figure 3.26 Cond(H) Over The Flight Profile

Due to the random nature of the search method, each optimization is not exactly repeatable, but the optimization becomes more repeatable as the density of points in the allowable region for alternate receiver points increases. If one is willing to wait long enough, the results can be made to be fairly repeatable. There is a question of whether it is necessary or desirable to have exact repeatability for this method. There is no guarantee that moving the receivers one by one to the positions that reduce the cost function the most will come close to achieving the elusive absolute minimum cost. Having some variability in where the receivers are placed each time (caused by a lower density of points) may give the method more robustness, at the cost of increasing the time it takes to get useable results (by requiring multiple runs for the same initial conditions). In this research, however, a middle ground is chosen. The alternate receiver point density is chosen high enough to obtain near repeatable (but not identical) results. This yields fairly consistent results, without losing all the robustness obtainable by evaluating multiple runs, yet does not alway require that multiple runs be done in order to have confidence in the results.

The resulting arrays are usually very similar, differing mainly in the exact placement of the receivers on the ground, rather than overall array shape/configuration. This consistency lends some confidence to the results.

The cond(H) traces are consistently improved by the use of this method. A good example of this improvement is shown in Fig. 3.26. Every run of this routine with some random initial array has improved the cond(H) trace over the transmitter flight profile, at least a little bit. If the initial array is very well chosen so that little improvement is possible, however, this routine will at worst leave the array unchanged. This means that the routine always works, at least to some extent, and is at least guaranteed not to make things worse. While no claim of global optimality can be made, it can be said that this optimization technique **always** either improves the array with respect to the flight profile or leaves it unchanged. It cannot diverge, by its very nature.

Symmetry:

The random search technique is not well behaved with respect to symmetry, in most cases. Although the random nature of the search did help to avoid getting stuck in small relative minima in the cost function, the very nature of the one-by-one movement of receivers tends to destroy symmetry. From the thought experiment with four receivers that resulted in the tetrahedron, it is normal to expect that a globally optimal array will have a degree of symmetry. The lack of symmetry in the optimized receiver arrays suggests that they are not truly 'optimal,' but simply better than they were before. This is not an unexpected problem, and does not prove to have a major negative impact on the design process. Much insight into good receiver array design can be gained simply by seeing the kinds of changes that improve the initial arrays.

Sensitivity to Initial Array Configuration:

The optimized arrays produced by the random search technique are not heavily dependent on their initial array configurations. Given enough iterations, the arrays generally converge to similar configurations and similar minimum costs, if the arrays have similar numbers of receivers and the cost functions use the same transmitter trajectory. It is not known whether this is a result of the 'optimality' of the solutions or just an artifact of this optimization method. Most likely, this method can take the cost down to within a neighborhood of the global minimum, but gets 'hung up' on some nearby local minimum because it can only move one receiver at a time. Although the optimized arrays tend to look similar, they are never exactly the same, even when the initial arrays and cost functions are identical. Possibly, if the random search uses enough points, the same results would be achieved every time for identical initial arrays, but the amount of time it would take to investigate this is prohibitive, so it was not done.

Since this technique takes an almost arbitrary initial array and improves it just about to the tradeoff between receiver spacing and aircraft altitude (for planar arrays), perhaps it could be used to obtain good initial conditions for more mathematically rigorous optimization techniques. Since those techniques tend to be sensitive to initial conditions, this idea seems plausible. Experimentation has shown this idea to have merit in circumstances where symmetry cannot be used to advantage, or where the cost function is so poorly behaved that initial array configuration totally determines whether the result is worthwhile. An instance of note is when the GDOP is used in the cost function instead of the condition number. The GDOP function does not lend itself towards global optimization by the numerical optimization routines provided in MATLAB, often 'hanging up' with no appreciable changes in the receiver array. Using the results of the Monte Carlo technique to initialize the MATLAB optimization programs resulted in better final arrays than were generally achieved with the input of arbitrary initial conditions.

Effectiveness:

The random search technique proves quite effective in optimizing receiver arrays with respect to a given cost function. Part of this effectiveness is due to the its inherent robustness. The technique is rather insensitive to initial array configurations, complicated transmitter trajectories, and constraints on the receiver locations (imposed by maps, signal strengths, etc.). Also, the very nature of the technique prevents it from making the array worse. The worst this method can do is to leave the array in its initial form. This robustness makes it likely that a satisfactory receiver array will be found, if it is possible, given the constraints on the problem. Whether a satisfactory array can be found or not depends mainly on the feasibility the cond(H) levels desired throughout the flight profile. A desired level of cond(H) may be lower than is possible for the given flight profile and receiver array. For example, if the profile is at too low an altitude, the number of receivers allowed in the array may not be enough to adequately cover the profile. This is really not a fault of the optimization routine; it is simply an illustration of the tradeoffs involved in receiver array design. Running the optimization routine a few times gives one a feel for what is possible and what is not, for a given number of receivers and flight profile, so this has not proved to be too much of a problem.

There have been some cases in which the random search routine has been especially effective, and especially ineffective. The random search technique is especially effective at improving receiver arrays given complicated flight profiles or stringent design constraints, such as a very low minimum altitude or a small number of receivers in the array. The random search technique does not use cost function gradient information, so it has no problems with asymmetric or complicated flight profiles. It just puts receivers wherever the locations lower the cost the most, so it will work to some extent even for difficult problems. This method is the one to use for a difficult or poorly behaved optimization problem. The unusually ineffective cases occur when the optimization parameters are poorly chosen. This optimization method requires the transmitter points to be spaced closely together and the alternate receiver points to be of sufficient density and chosen over a large enough region (at first) to ensure the method's convergence on a good result. Otherwise, although the method won't make the array worse with respect to the specified profile, it might not improve things very much, and may make the array totally useless for anything else (thus wasting the designer's time). The generation of 'optimally stupid' results is the main hazard of optimizing a receiver array with respect to one flight profile, and can frequently happen with this technique if the optimization parameters are poorly chosen. Therefore, if the routine seems of little use on some problem, the first things to check are the optimization parameters.

Summary:

The random search method of receiver array optimization has proven to be quite useful in this research due to its relative insensitivity to initial conditions, its ability to avoid 'hangups' on relative minima early in the optimization, its consistency, and its value as a tool to find good initial conditions for using more rigorous numerical optimization routines. With the parameters properly chosen, this optimization routine is an effective tool for the optimization of receiver arrays, especially for difficult cases. Unfortunately, there is no indication that the arrays optimized by this technique are in fact the best possible. Therefore, this tool should not be used (by itself) if it is desired to find the elusive globally optimum array for some flight profile. In that case, its use should be limited to finding good initial conditions for a more rigorous optimization routine or being a check on the results. (If the random search finds a better array than the supposed optimal one by moving one of its receivers around, then the array in question is not really optimal.)

3.5.4 SQP Optimization Routine: **constr.m.** The MATLAB optimization package includes several numerical optimization routines. One of these, constr.m, has proven to be useful for receiver array optimization. The optimization routine performs constrained optimization on a vector of independent variables with respect to a user defined objective (or cost) function. The routine uses a sequential quadratic programming (SQP) method to search for the minimum cost. This method has proven to be quite successful in optimizing the receiver arrays, given good initial conditions and certain kinds of flight profiles.

The constr.m routine is designed to take a cost function and a vector of initial values for the independent variables, and find the values of the independent variables that minimize the cost within some neighborhood. By itself, it is a generic optimization tool. A program had to be constructed around constr.m in order to optimize receiver arrays. The program works as follows:

1. **The initial receiver array is formed.** It could be hand picked, randomly generated, data from previous experiments, etc.. The initial receiver array consists of the initial conditions for the independent variables used in the constr.m routine. The initial list of receiver positions is decomposed into one long vector, the initial conditions vector to be used in the constr.m routine.
2. **The flight profile for the transmitter is formed.** The flight profile is used to form the cost function for the optimization. As this is a numerical optimization, the flight profile is discretized into points, as if there were transmitters at many points simultaneously along the desired flight profile. It is the mean or maximum condition number of all the H matrices formed at these transmitter points that becomes the cost function for the optimization.
3. **constr.m is run.** The initial conditions vector and the transmitter flight profile points are fed into the constr.m routine, which then finds the vector of independent variables that minimizes the cost within some neighborhood.
4. **The receiver array is reformed.** The vector of independent variables that minimize the cost function is taken from constr.m and reformed into an array of receiver positions.

5. The results are plotted. The receiver array is plotted, along with contours of cond(H), to show the array's performance.

The utility of this technique can be evaluated by considering the four criteria of consistency, symmetry, sensitivity to initial array configuration, and effectiveness, as is done for the Monte Carlo technique.

Consistency:

The optimization routine based on constr.m proves to be fairly consistent. Experiments are always repeatable for a given set of initial receiver locations and transmitter points. Optimized arrays tend to have similar characteristics, if their initial array configurations or their flight profiles are similar. This method consistently improves the array with respect to the given cost function, at least somewhat. The only way in which this method is not consistent is that using different initial conditions or flight profiles may totally change the resulting array. Even so, this method is generally more consistent than the Monte Carlo technique.

Symmetry:

This method is quite well behaved with respect to symmetry. Symmetric initial conditions and symmetric flight profiles generally result in symmetric optimized arrays. As the lowest cond(H) at a single point is obtained by a tetrahedral array, which has a high degree of symmetry, it stands to reason that the optimum array for two or more transmitter points arranged in a symmetric pattern should be symmetric as well. The fact that symmetry is preserved for the most part using this method implies that this technique has the potential for producing arrays that are close to being globally optimal, given the proper initial conditions.

If the initial array is not symmetric, some degree of symmetry is imposed on the array through the optimization process. Perfect symmetry is not restored, not surprisingly, but generally the resulting array has some degree of symmetry, even if the initial array does not. Of course, this effect can only work with those receivers that are allowed to move by the topography of the cost function's (many dimensional) 'surface'. It has been found that if the receivers are too close to the horizon throughout the entire flight profile, they are not moved very much by the optimization process. Therefore, those receivers will keep whatever configuration they had in the initial array, symmetric or not. These results indicate that symmetry is indeed an important design parameter, as expected.

Sensitivity to Initial Array Configuration:

Unfortunately, the constr.m routine is very sensitive to initial array configuration. The initial receiver array that is fed into this routine determines what results are obtained. The cost function for a planar array simply has too many dips and peaks in it for a global optimization to be consistently found, using gradient information. The fewer dips and peaks there are in the cost function, the less sensitive this routine is to initial array configuration. With a single transmitter point and a small number of receivers (four to six) in the array, this routine finds solutions that look very good. With only four receivers and one transmitter point, this routine finds the tetrahedron in unconstrained optimization, and produces a consistent trigonal pyramid shape in the optimization constrained to the plane, for different initial conditions. Unfortunately, when the number of receivers is increased to more than ten or so, the optimization often gets hung up on local minima of the cost function that obviously are not the best possible. Sometimes one can hand pick better arrays, sometimes. Since there need to be more than ten receivers on the ground for a good, large planar array, the placement of initial conditions is critical.

The initial array configuration problem can be overcome in three ways: running multiple runs of the same flight profile with different initial arrays, hand picking the initial arrays using general guidelines for good array design, or running the Monte Carlo optimization routine first on some arbitrary initial receiver array to refine it for the constr.m optimization. Each of these methods has its merits, and, in a search for the global optimum, it would be wise to do all three and compare the results.

Once the results are in, they can be put as initial conditions into the Monte Carlo optimizer (which does not use gradient information) to see if an even better array can be found. Since the Monte Carlo optimizer does not use gradient information, it may be that the result of this hybrid optimization could be run once more through the constrained optimization (constr.m) to fine tune it even more. After that many iterations, it is likely that an array very close to the global optimum could be found.

Effectiveness:

The effectiveness of this optimization technique varies. When the initial arrays are properly (luckily?) chosen, this routine is very effective, as shown in Fig. 3.27. With an unfortunate choice of initial array, this optimization is not very good at all, as shown in Fig. 3.28. Sometimes, the technique does not even converge, wasting time, and forcing a change in flight profile, initial array, or tolerance parameters such as maximum number of iterations. Since the progress of the optimizer can be monitored from time to time on the screen, one can sometimes tell when it is not working well, and 'kill' it early, so it can be rerun without wasting as much time.

When the initial arrays are properly chosen, however, the optimizer can do a very good job. When it works, its results have more weight than the results of the Monte Carlo optimizer, since the constr.m optimizer is guaranteed to find a minimum, local or otherwise, but there are no guarantees that the Monte Carlo optimizer will find even that. This routine is also quite good at fine-tuning a hand picked array or one optimized by the Monte Carlo method. It will not move the receivers around much, most likely, but the cond(H) profile will show improvement, often being smoother or having spikes that are not as large.

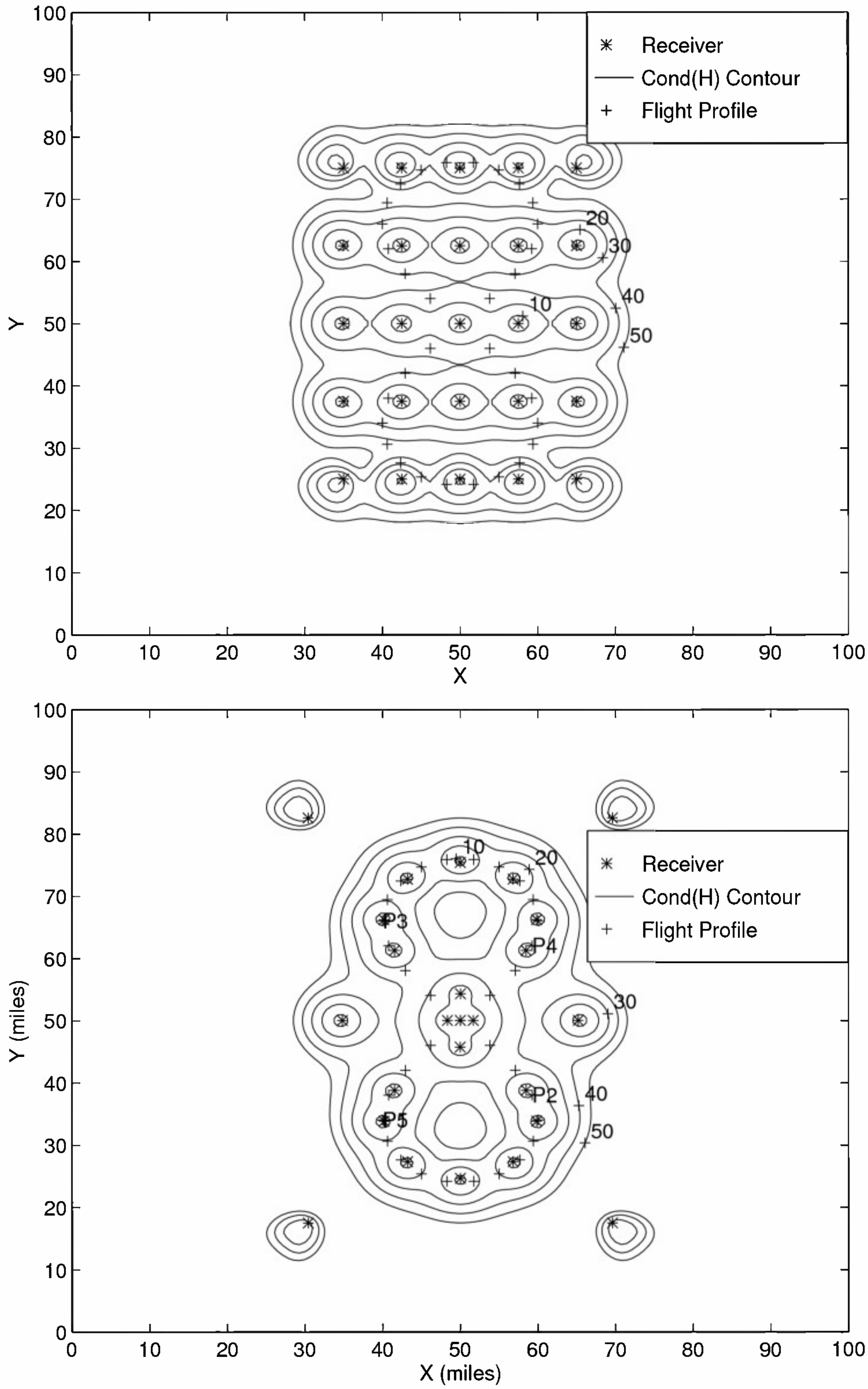


Figure 3.27 Initial and Final Arrays – Good Run

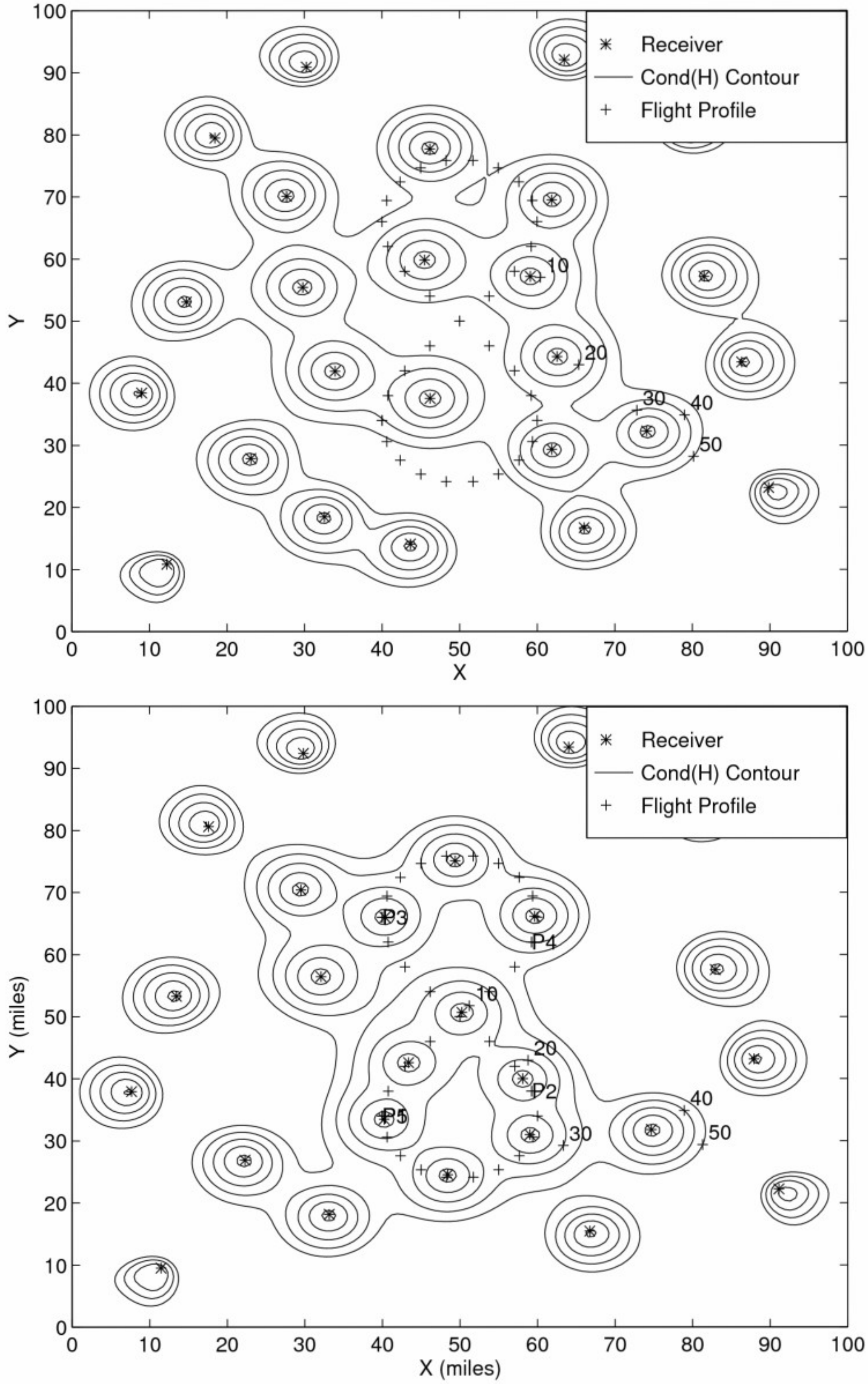


Figure 3.28 Initial and Final Arrays – Constr.m Got 'Hung Up'

Summary:

The array optimizer based on constr.m can be capricious (due to the cost function), but it is worth using it for the chance of obtaining the really good results that are possible when all the conditions are right. Although there is high sensitivity to initial receiver array configuration, every experiment is repeatable, and the results are at least guaranteed to be at some minimum in the cost function. This optimizer is the right one to use for fine tuning an array, or in a search for the elusive globally optimum array for some specific flight profile.

IV. Results and Analysis: Receiver Array Optimization

This chapter shows the results of the numerical optimization research. The experiments fall into two main categories: the optimization of receiver arrays with respect to a single fixed transmitter design point, and the optimization of receiver arrays with respect to multiple transmitter points (discretized trajectories). The results of the optimizations with respect to the single transmitter point are discussed first. The array optimization problem is simplified due to the presence of only a single transmitter point in the cost function. The optimization program based on constr.m is modified to take advantage of the simplicity. This modified program is used to solve the array optimization problem for a single transmitter point, providing much insight into good array design. Next, the more complex multiple transmitter point optimization problem is discussed. To avoid getting bogged down in excessive details, only three representative flight profiles are considered for the transmitter profiles: a circle, a figure eight, and a grid pattern of points. Several initial receiver arrays are tried, for each profile, and the results of the optimizations are shown and discussed. Once the results of the array optimizers for the three flight profiles are discussed, the knowledge gained can be combined with the insights from the single point optimization to develop a set of guidelines to aid in the quick and effective optimization of planar arrays with respect to some arbitrary flight profile, and vice versa.

4.1 Optimization With Respect to a Single Transmitter Location

The first flight profile considered for the transmitter is the most elementary one: a single, fixed point. This case is especially easy to set up and analyze due to its simplicity. Although a simple case, much can be learned about good receiver array design by finding the receiver array configurations that produce the lowest condition number for this case, and looking at how the results can be applied to array design using more complicated flight profiles.

As was shown in the thought experiment earlier in this chapter, the receiver array optimization problem with respect to a single, fixed transmitter location can be posed in a simplified form that is most insightful. Having only a single transmitter point means that there is only one H matrix under consideration, which allows the angular position terms in the H matrix to be reduced to simply position terms (on the unit sphere). This is done by recognizing that the projection of any array of receiver locations along receiver-transmitter line of sight onto the unit sphere centered on the transmitter yields a new receiver array that will have the same H matrix. Only the angular orientation of the receivers from the transmitter's point of view affects the H matrix. This new receiver array can be optimized much easier than the old one, and can be re-projected out onto some surface (like the ground plane, or some other surface of allowable receiver locations) when the optimization is complete. Now that the receivers are all on the unit sphere, the angle component cosines that form the elements of the first three columns of H can be reduced to simply position components of the receivers' location, as shown for the four receiver case in the equation below.

$$H = \begin{bmatrix} \cos\theta_{1x} & \cos\theta_{1y} & \cos\theta_{1z} & \cos\theta_{1t} \\ \cos\theta_{2x} & \cos\theta_{2y} & \cos\theta_{2z} & \cos\theta_{2t} \\ \cos\theta_{3x} & \cos\theta_{3y} & \cos\theta_{3z} & \cos\theta_{3t} \\ \cos\theta_{4x} & \cos\theta_{4y} & \cos\theta_{4z} & \cos\theta_{4t} \end{bmatrix} = \begin{bmatrix} X_1 & Y_1 & Z_1 & 1 \\ X_2 & Y_2 & Z_2 & 1 \\ X_3 & Y_3 & Z_3 & 1 \\ X_4 & Y_4 & Z_4 & 1 \end{bmatrix} \tag{4.1}$$

With the H matrix now written as a simple function of the receivers' positions, the optimization process can be easily performed. Two main cases were tested for this simple problem: the fully unconstrained case and the case where the receivers were constrained to all lie in one plane. Both cases were handled well by the constrained optimization routine (constr.m). It performs so well on this simplified problem that the Monte-Carlo search method was not needed. However, the program that sets up the constr.m routine did need to be changed to take advantage of the simplified problem. This modified program works as follows:

1. **The initial receiver array is formed.** The initial receiver arrays are formed, either by random generation or by hand-picking.
2. **The array is projected onto the unit sphere.** The initial receiver arrays are projected onto the unit sphere, forming the initial conditions for the independent variables in the optimizations.
3. **The cost function is formed.** The cost function is formed. For this case, the cost function is simply the condition number of the H matrix formed using the receiver positions on the unit sphere.
4. **const.m is run.** The constr.m routine is run, which finds a vector of independent variables that minimize its objective function, subject to constraints. For this problem, the independent variables are the receivers' position components along the coordinate axes, and the cost function is the condition number of the H matrix. For the unconstrained optimization, the only 'constraints' on the variables are that the receivers must lie on the unit sphere, which just ensures that a valid H matrix is found. For the optimization in which the receivers must stay in one plane, additional constraints are needed to force the receivers to be able to be projected back onto the plane after the optimization. These constraints are very simple: the receivers' positions are constrained to be in one half of the unit sphere, i.e., the $z \leq 0$ hemisphere.
5. **The results are plotted.** The resulting arrays are taken from constr.m and plotted with cond(H) contours to see their performance. It has been found that it makes sense to keep the receivers in spherical array format, rather than to project them back out onto some surface of desired receiver locations (like the ground plane). For a spherical surface, this is so because such a projection

would be identical in all but scaling, so there's no reason to do it. For a planar array, the projections could be useful but are problematic because there are always a few receivers on the horizon that mess up the projection. In either case, it is possible to analyze the array on the unit sphere just as well, so nothing has been lost.

When looking at the optimized receiver array shapes, a word of caution is in order. The receivers can be moved along their line of sight vectors to and from the transmitter without affecting the condition number. The array shapes shown in the following results are not the only ones possible. Changing the magnitudes of the LOS vectors can significantly distort the way the arrays look, but will not change the condition number of H near the transmitter design point. This can be a useful property, and should not be ignored. Although the actual optimization of the receiver arrays proceeds best with the receivers constrained to lie on a sphere, an array projected back onto a different surface (like a plane or ellipsoid) may be better for a particular application.

4.1.1 Results: Unconstrained Optimization. Many different initial receiver arrays were run through the unconstrained optimization routine. The number of receivers ran from just four to 25, and many different initial array configurations were tried. Regardless of the number of receivers in the array, the results of the optimization were all quite similar. The results shared two things: symmetry and minimum condition number. The optimized receiver arrays had some degree of symmetry (at least for small numbers of receivers), and all the optimized arrays produced H matrices whose condition numbers were around $\sqrt{3}$.

It turns out that the optimal arrays have some degree of symmetry, at least if the numbers of receiver in the arrays are small. This symmetry is readily apparent with arrays of few receivers. Figs. 4.1 and 4.2 show the optimal configurations of a four receiver array and a five receiver array. The four receiver array is configured in a tetrahedron, with the origin at its center. The five receiver array is configured in a

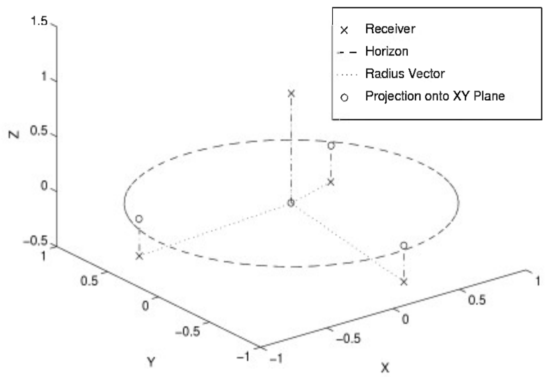


Figure 4.1 Optimal Array, 4 Receivers

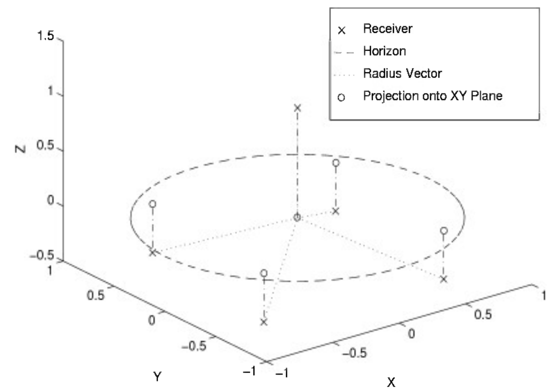


Figure 4.2 Optimal Array, 5 Receivers

square pyramid. For these cases with low numbers of receivers, the array shape and symmetry are apparent. It is much more difficult to find the shape and symmetry of an array with a large number of receivers, such as the array shown in Fig. 4.3, though the symmetry may indeed exist. This difficulty may be caused by the non-uniqueness of the optimal receiver array configurations, given large numbers of receivers. A good example of this is that of Fig. 4.4, an optimal eight receiver array formed out of two optimal four receiver arrays (two tetrahedra). It is difficult to see the overall symmetry of this array, and could be impossible if the tetrahedra were of widely different orientations (or if the five receiver optimal array were used instead of one of the tetrahedra). The 25 receiver array shown in Fig. 4.3 could well be formed out of five tetrahedra and a square pyramid, five square pyramids, or some more complicated combination of optimal arrays, just so long as the number of receivers in the array is 25. Therefore, no attempt is made to describe optimal shapes or configurations for arrays with large numbers of receivers. Many shapes are possible. In large arrays, the symmetry is at the component array level, not over the entire array.

In addition to a degree of symmetry, all the optimized arrays produced H matrices that had condition numbers close to (within 6 percent of) $\sqrt{3}$. This supports the results of the thought experiment conducted earlier this chapter. The lowest

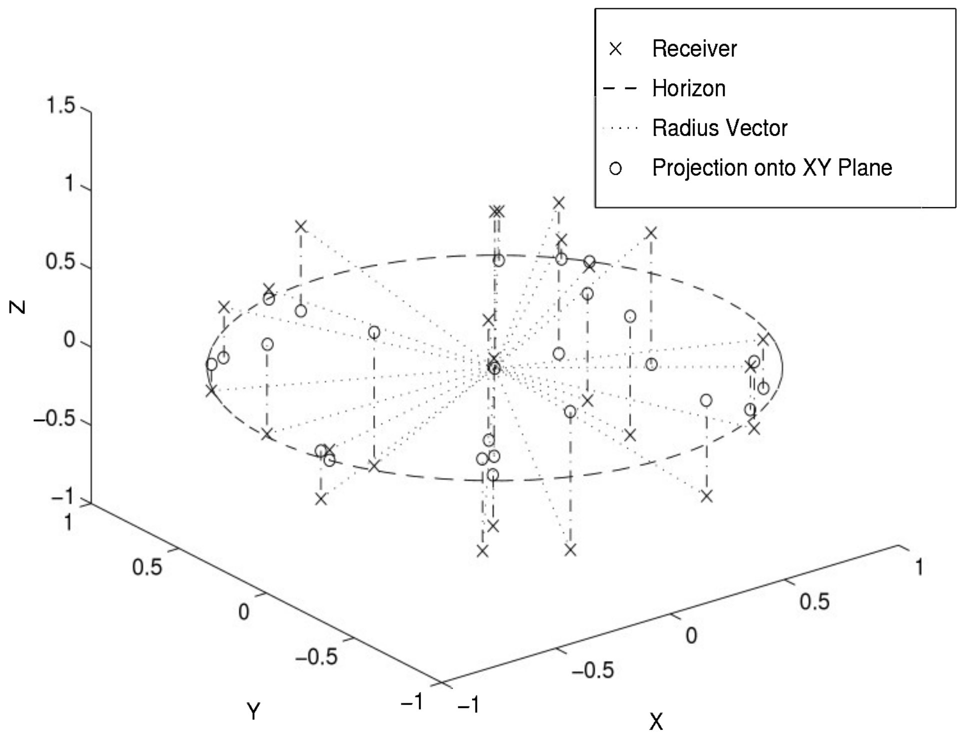


Figure 4.3 25 Receiver Array, Optimized About the Origin

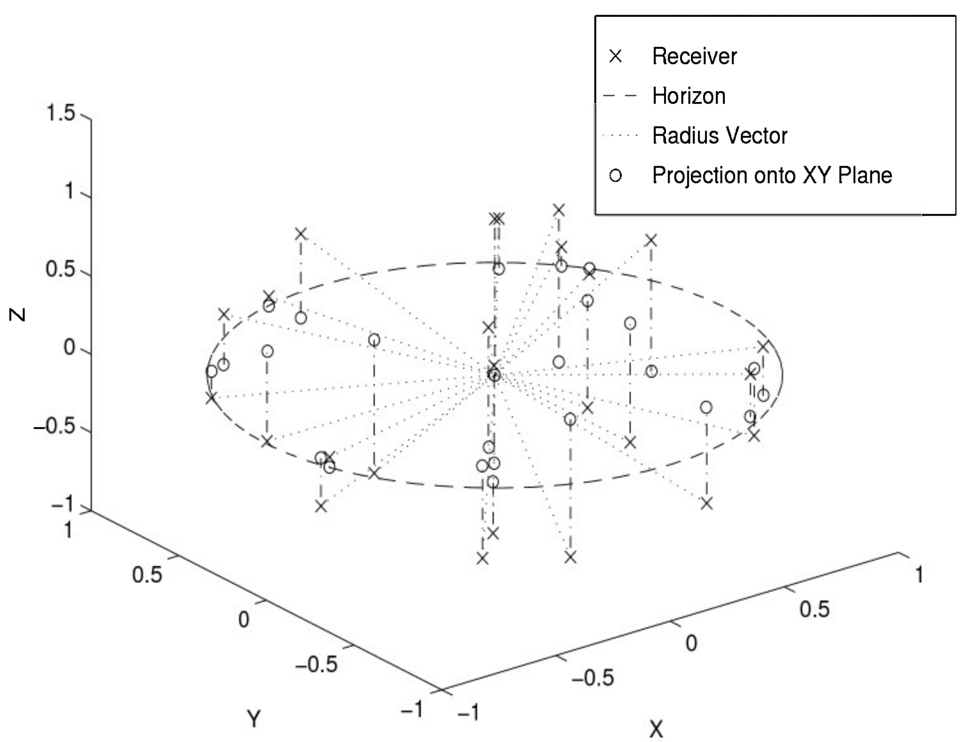


Figure 4.4 8 Receiver Array (Two Tetrahedra)

possible condition number for the H matrix is $\sqrt{3}$. Since the optimized arrays produced H matrices which had the minimum condition number, it is reasonable to believe that these optimized arrays are in fact globally optimum with respect to cond(H).

Design Insights:

Several array design insights are found from looking at the unconstrained optimization results:

1. **Cond(H) is independent of the number of receivers.** The minimum condition number can be obtained with as few or as many receivers desired, as long as there are at least four receivers in the array. Therefore, for this case of unconstrained receiver positions, the number of receivers can be decided entirely by other design issues than the cond(H), (such as filtering considerations or logistics).

2. **The array should be re-designed when changing the number of receivers.** Given an already optimal array, if it is desired to change the number of receivers in the array, it is better to re-design the array using the new number of receivers than to simply add a receiver somewhere. In general, simply adding a receiver somewhere to an optimal array will raise the condition number somewhat. However, adding two optimal arrays together does produce a composite array that is optimal also. The thing to keep in mind is that receivers should not be added or subtracted from these arrays without some thought, if low cond(H) is desired. Unlike the GDOP metric, the cond(H) may actually increase with the addition of another receiver to the array, if the array is not redesigned as well.

3. **The farther the receivers are from the transmitter, the better.** To see this, look at Fig. 4.5, which shows cond(H) as a function of transmitter position around a fixed tetrahedral array. The cond(H) is at a minimum at the array's center (the transmitter design point), and grows larger as the

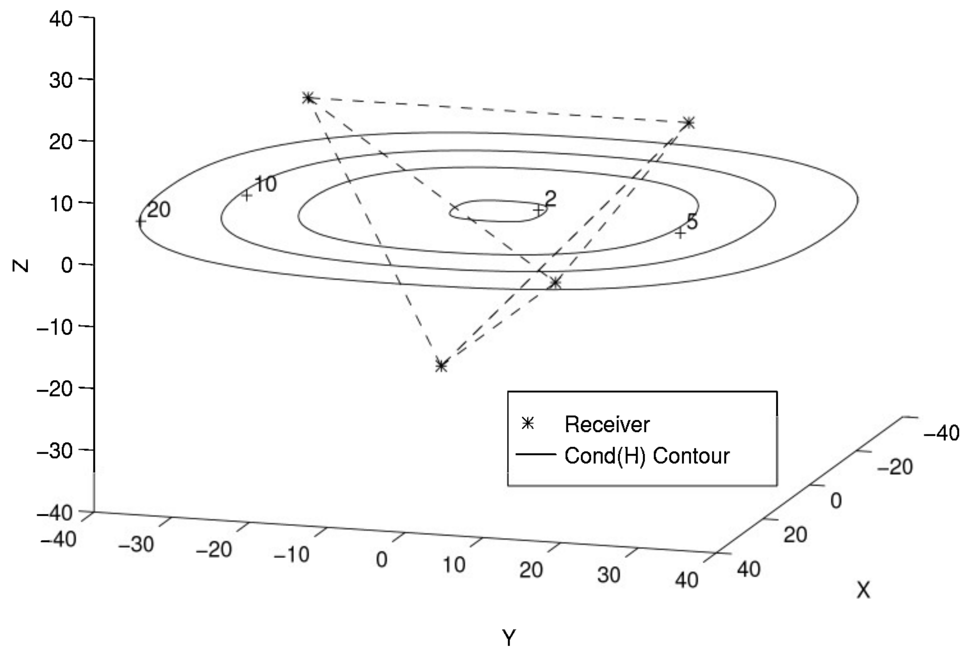


Figure 4.5 4 Receiver Array, Optimized About the Origin

transmitter moves away from the center. Once the transmitter goes outside the array to any appreciable extent, the condition number grows large. Therefore, if low cond(H) is desired, the transmitter must remain within the array. If the application allows it, there is an easy way to ensure that the transmitter remains inside the receiver array: simply make the receiver array so big that it encompasses the entire region in which the transmitter will move. This will enlarge the region of low cond(H) until it fills the entire region through which the transmitter moves.

4.1.2 Results: Planar Array Optimization. Many different initial arrays were run through the optimization, just like for the unconstrained case. This constrained optimization was a bit more sensitive to initial array configuration than the unconstrained case, so more runs were done with different initial arrays to find the best planar arrays for various numbers of receivers. The results of the planar array optimization were even more alike than those of the unconstrained optimization. The optimized arrays all had the same overall configuration, and produced H matrices that had similar condition numbers. Although the minimum cond(H) produced was higher than the minimum possible, it was not that much larger, indicating that a planar array is really not that much worse than a fully three dimensional array, at least with respect to a single, fixed transmitter point. Unfortunately though, the low cond(H) region around the transmitter design point of a planar array is much smaller than that of a three dimensional array. This heavily impacts the design of planar arrays in cases of appreciable transmitter motion (such as the SARS).

The optimized planar arrays all were of similar configuration. Figs. 4.6 and 4.7 show the basic configuration that was shared by all the arrays. Although the receiver arrays, were designed on the $z \leq 0$ hemisphere as discussed earlier, these figures show the receivers projected back onto the ground plane for clarity. To correctly interpret these plots, it must be understood that the transmitter is located a small distance directly above the origin (above the plane of the array). The transmitter location is not shown in the figures because it was found to diminish the figures' clarity.

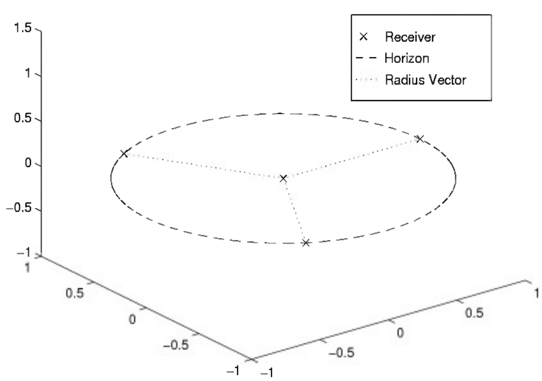


Figure 4.6 Optimal Planar Array, 4 Receivers

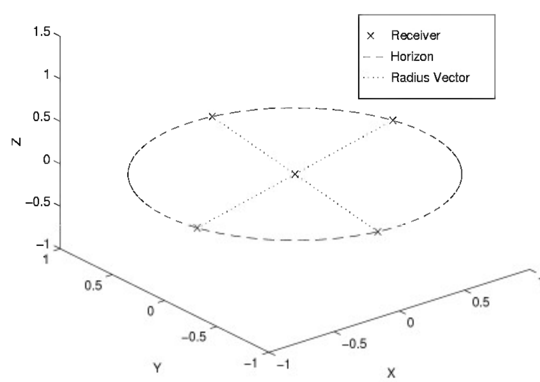


Figure 4.7 Optimal Planar Array, 6 Receivers

As can be observed, the best configuration is to have at least one receiver directly below the transmitter (the receiver(s) at the origin), and have the rest out along the horizon. For a four receiver array, three receivers are located on the horizon, spaced at 120 degree increments, and the fourth receiver is directly below the transmitter. With five receivers, the configuration is identical, except that there are two receivers on the ground directly underneath the transmitter. For six receivers, the best configuration is four receivers out on the horizon, as shown in Fig. 4.7, with the remaining two directly underneath the transmitter. As the number of receivers grows, this same kind of configuration is found. It is surprising that having two or more receivers in exactly the same location can improve the condition number, albeit a small amount. However, the improvement in cond(]itH) is small enough that such duplicate receiver positions can be ignored, saving the extra receiver(s) for a better location with respect to other design criteria like line of sight, or carrier phase ambiguity estimation.

As with the unconstrained optimizations, symmetry is apparent in arrays with small numbers of receivers, but is lost when the number of receivers gets large. The configuration stays the same, however, as is shown by Fig. 4.8: The lowest cond(H) is produced when most of the receivers are placed all around the transmitter out on the horizon and the rest are all lumped together at one point underneath the transmitter. The difference in cond(H) produced by adding a receiver to the ring on the horizon or adding one to the pile directly underneath the transmitter is negligible.

The cond(H) produced by these arrays ranged from 2.41 to 2.5. This is not a significant variation in cond(H). The minimum cond(H) of 2.41 is slightly higher than the 1.732 global cond(H) minimum, but not very much so. This indicates that a planar array is potentially not much worse than a fully three dimensional array, at least if the transmitter's motion is limited to near the design point.

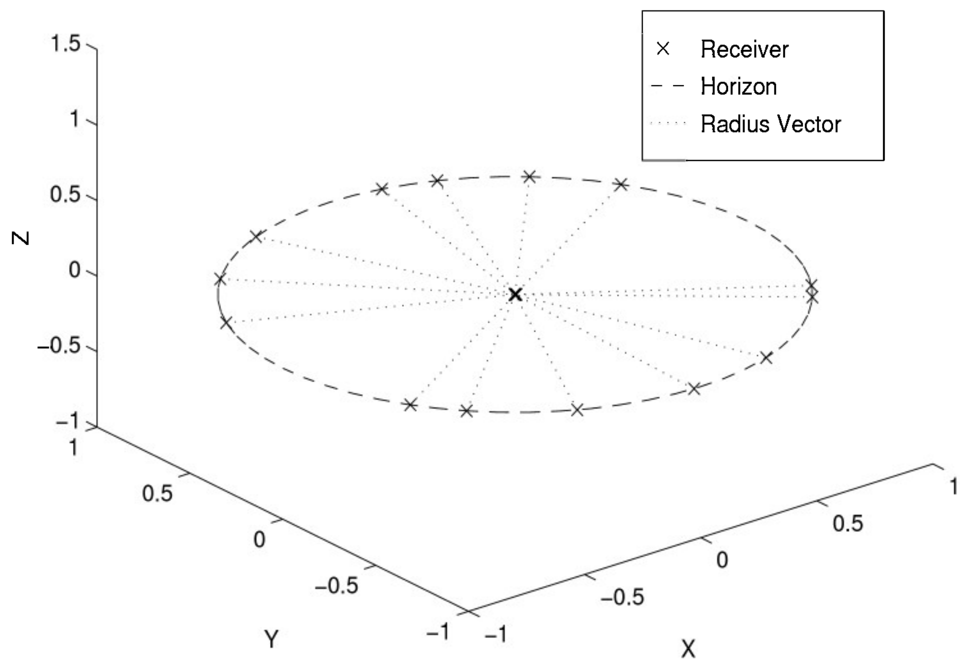


Figure 4.8 Optimal Planar Array, 25 Receivers

Unfortunately, limiting the transmitter to the region close to the design point proves much more restrictive in planar arrays. This problem is illustrated by Fig. 4.9, a contour plot of the cond(H) fields at one mile altitude over a four receiver planar array. This array is approximated from the optimal planar array shown in Fig. 4.6. The receivers set in the circle of thirty mile radius represent receivers on the horizon, and the receiver at the origin represents the receiver that needs to be directly below the transmitter. The transmitter design point is at one mile directly above the origin. This is indeed an approximation of the optimal four receiver planar array, but it does produce a low cond(H) of 2.6 at its design point, so it is a good approximation. As can be seen, although the cond(H) is still quite low at the design point, the cond(H) increases rapidly as the transmitter moves horizontally away from its design point. Instead of being able to move over the entire width of the array (at that altitude) and still obtain a cond(H) less than five, as is the case with the tetrahedron of Fig. 4.5, the transmitter must stay within a two mile circle of

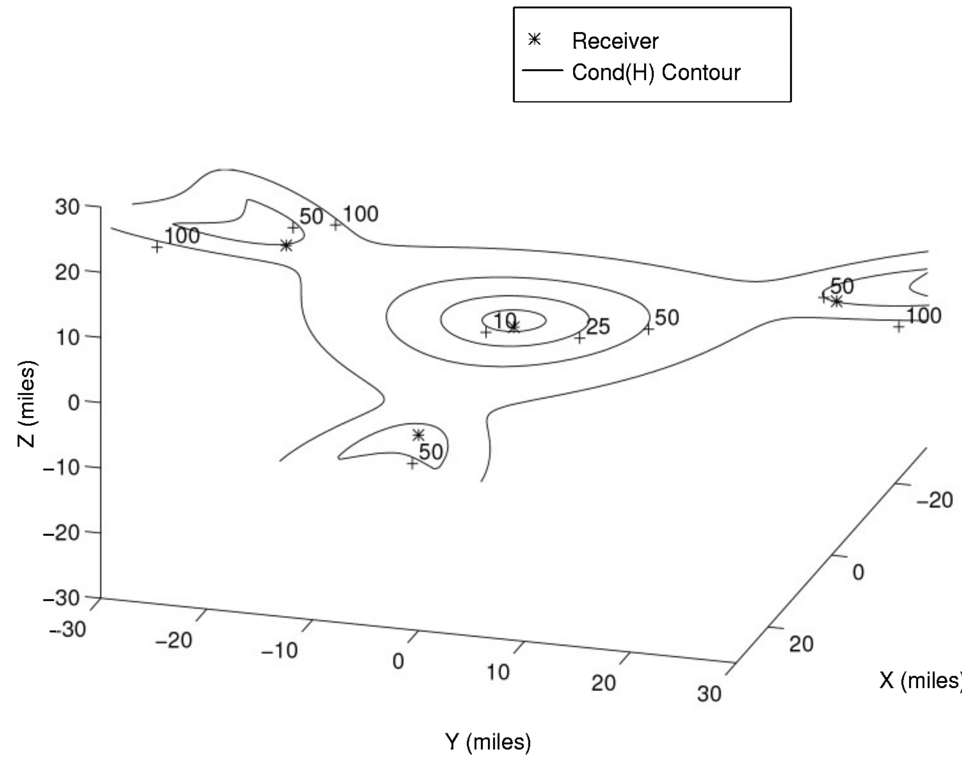


Figure 4.9 Optimized 4 Receiver Planar Array, Cond(H) Fields Plot

the center receiver. Therefore, although the cond(H) values are very similar at the design point, the tetrahedral array allows much more receiver movement for a given array size than the planar array does.

Design Insights:

Several planar array design insights are found from looking at the optimization results:

1. **Cond(H) is independent of the number of receivers in the array, as long as the transmitter does not move.** The cond(H) stays around 2.45, regardless of the number of receivers in the array. Of course, Fig. 4.9 shows that if it is desired for the transmitter to move around a lot, then more receivers are needed in order to make the low cond(H) region large enough to be satisfactory.

2. **The only receivers that matter in a planar array are ones that are directly below the transmitter or on the horizon.** Adding receivers anywhere else does not improve the cond(H). This can be used to advantage in planar array design: There is no need to put receivers anywhere other than directly underneath the flight trajectory and at points out along the horizon. Also: there need only be three or four receivers out on the edge of the array (the horizon), as long as the transmitter doesn't get too close to them, or they run the risk of losing line of sight due to the terrain.

3. **The best planar array performance is at the maximum altitudes the test aircraft can attain.** Like the three dimensional arrays, planar arrays can be expanded in all dimensions to increase the size of the low cond(H) region about the transmitter design point. Unfortunately, the receiver that needs to be the farthest from the aircraft is the one directly underneath the transmitter. The farther this receiver is from the transmitter, the larger the array's low cond(H) region will be. The only way to increase the distance between this receiver and the transmitter is to fly the aircraft at higher altitudes. Therefore, the transmitter altitude is a **critical** parameter of planar array design. It should be as high as possible.

4. **A planar array can be significantly improved by flying a receiver high above the test aircraft over the array.** A fixed planar array cannot achieve the low cond(H) of the tetrahedral array, and its region of low cond(H) is very small, comparatively. This problem can be eliminated if the receiver directly below the transmitter could be made to move along with the transmitter, always remaining directly below the transmitter throughout the transmitter's flight. Of course, no receiver on the ground would be able to keep up with an aircraft, but the same geometry is obtained when a receiver is mounted onboard a second aircraft and flown directly above the test aircraft. This configuration would cut the required number of receivers for low cond(H) down to five or

six. The position of the flying receiver would need to be found, but this could be found by mounting a transmitter on the same aircraft as the one with the receiver. The pseudorange equations could be simultaneously solved to obtain the positions of both the moving receiver on the second aircraft and the original transmitter on the first aircraft. This configuration allows high accuracy to be obtained in low altitude flights, something not obtainable by a fixed, ground based array. Although this design entails using a receiver on a balloon, aircraft, or satellite as well as the receivers on the ground, the performance of the array is so much improved that this technique should be considered if high accuracy is desired.

4.2 Optimization With Respect to Multiple Transmitter Locations

Now that basic design insights have been gained from the results of the single transmitter point optimizations, the more complicated planar receiver array optimization problem with multiple transmitter points is attacked. Using the Monte Carlo search program and the program based on constr.m that were developed in this research, the receiver array optimization problem is numerically solved with respect to several representative flight profiles and initial arrays. While no global solution for all cases is found, the arrays optimized for these cases, representative of many found over the course of this research, show how to design good, if not globally optimum receiver arrays for practical use with flight testing. This discussion looks at the difficult problem of planar receiver array optimization. It applies what has been learned over the course of this research, and shows the kinds of array configurations that could be used to minimize the geometric sensitivity of ground based GPS-like systems like the SARS.

4.2.1 Initial Arrays and Transmitter Profiles. During the course of this research, a wide variety of initial receiver arrays and flight profiles were used in the optimization programs, with a correspondingly wide variety of results. The choice of

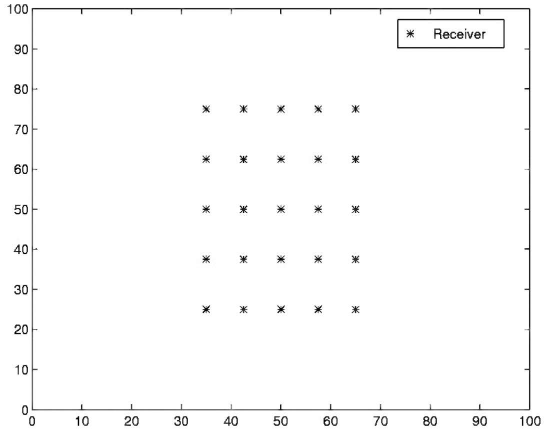


Figure 4.10 Initial Array: Grid

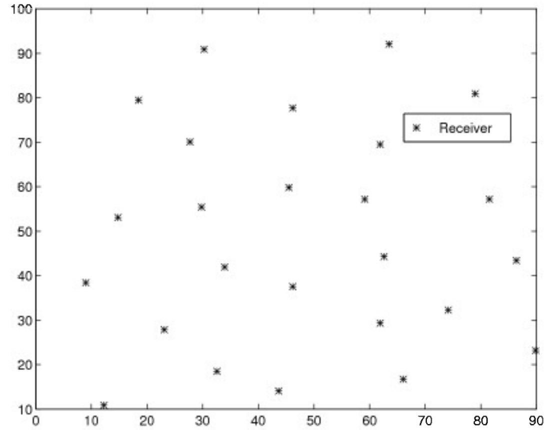


Figure 4.11 Initial Array: Irregular

initial array configurations and transmitter flight profiles can have profound effects on the outcomes of the optimization programs. Although it is not possible to show all the cases tried, the results fell into several general categories, which can be adequately covered by several representative initial arrays and flight profiles. Three initial arrays and three flight profiles are shown. The three initial receiver arrays are a receiver grid, an irregular, hand-picked receiver array, and an array that is the result of a previous Monte-Carlo array optimization. These initial arrays are shown in Figs. 4.10-4.12.

The three flight profiles are a circle, a figure eight, and a grid pattern that fills an entire region. The three initial arrays are so chosen because they reflect the three main kinds of initial array that have commonly been fed into the optimization routine: a symmetric, orderly pattern, an irregular mass of points, and an array that is already pretty good but may be improved. The flight profiles were chosen to reflect either typical test profiles, as were the circle and figure eight, or to generate a region of low cond(H) that could be used for any arbitrary profile, as was the case for the grid. The flight profiles are shown in Figs. 4.13-4.15.

All the flight profiles are designed at a constant altitude of one mile. Although this is artificial, it makes sense for the optimization process. As was found by plotting the cond(H) fields around a receiver array, the cond(H) becomes better behaved in

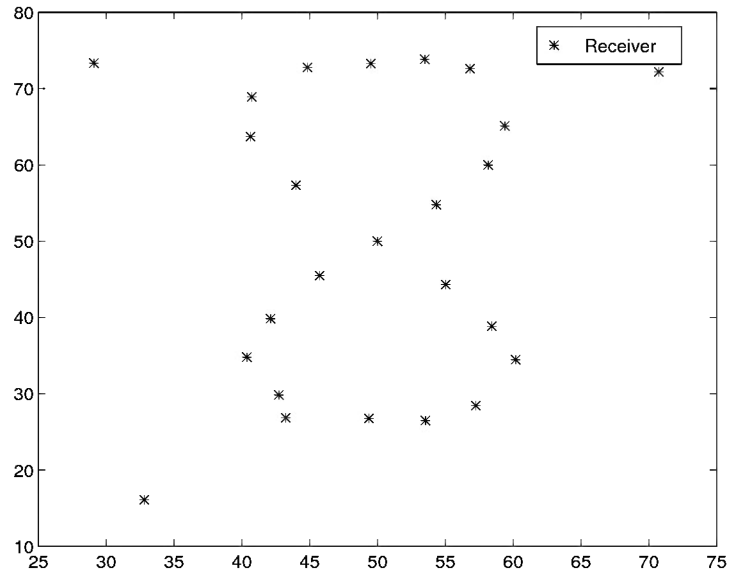


Figure 4.12 Initial Array: M.C. Optimized for Figure Eight

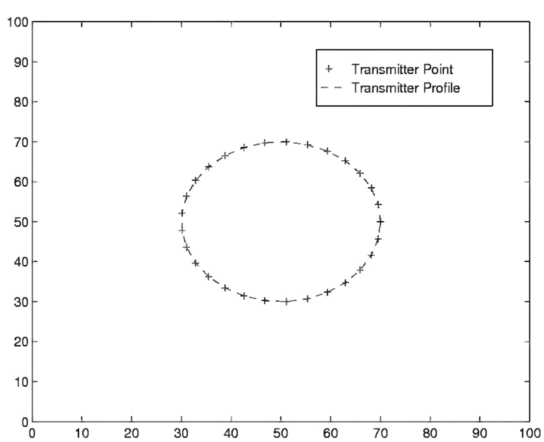


Figure 4.13 Transmitter Profile: Circle

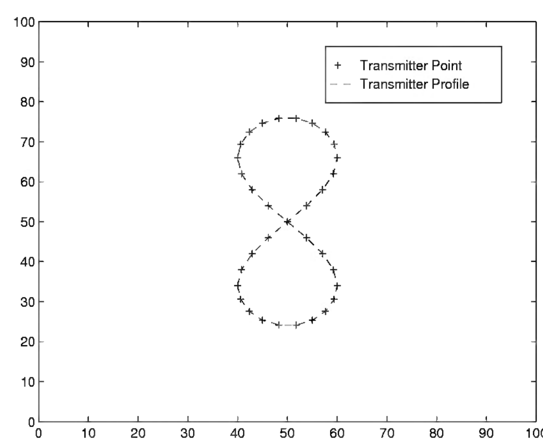


Figure 4.14 Transmitter Profile: Figure Eight

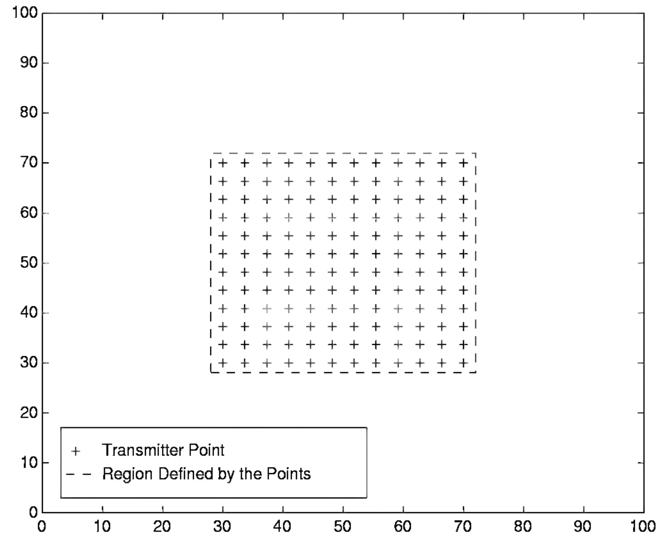


Figure 4.15 Transmitter Profile: Grid Pattern

a planar array the higher the aircraft can get (since the best areas occur above aircraft altitude limits). Therefore, *only the minimum altitude needs to be included in the design.* This altitude of one mile above the ground has been found to be a reasonable lower limit for flight profiles, based on a maximum of 25 receivers in the array, a desired cond(H) between ten and twenty, and a profile that allows at least twenty minutes or so of flight time (going 300mph) over the array. Lowering the minimum altitude would either cause the number of needed receivers to increase, force acceptance of a larger cond(H), or decrease the extent of the array. Doubling the density of receivers on the ground allows the aircraft's minimum altitude to come down by a factor of two, but as one can see, this quickly leads to point of diminishing returns. Doubling (or squaring) the number of receivers to get the aircraft minimum altitude down 2500 feet does not seem worth it. It seems better to just make the aircraft fly at least one mile up above the array. If high accuracy at low altitudes is desired, additional techniques need to be used to get the cond(H) down to reasonable limits without ridiculous numbers of receivers.

4.2.2 Results of the Array Optimization. The results are generated by pairing an initial array with a flight profile, feeding them into the optimization programs, then running the programs. Given the two optimization programs, the three initial arrays, and the three flight profiles, a total of eighteen optimized arrays are formed. All of these need not be shown, as some cases are duplicates or are not very useful. Of all eighteen cases, the ones of note are:

1. The three arrays produced by the Monte-Carlo optimizer (one for each flight profile). The Monte-Carlo optimizer proves unaffected by initial array configuration, so only one M.C. optimized array per flight profile needs to be shown. The others are duplicates. These three cases show the kind of results typical to the M.C. optimizer.
2. The arrays found by using the grid and irregular initial arrays with the constr.m optimizer for all three flight profiles. The constr.m optimizer is sensitive to initial array configuration, and the shape and number of points in the flight profile. These six cases illustrate the kinds of results typical to the constr.m optimization program.
3. The array found from using the constr.m optimizer on the results of the M.C. optimization with respect to the figure eight profile. This case shows how the constr.m optimizer can take a pretty good array and 'tune' it to get improved performance.

The resulting optimized arrays are shown in the following figures. The figures are organized with respect to flight profile. For each flight profile, the two initial arrays are shown superimposed on the profile, the arrays optimized by the constr.m and M.C. optimization programs are shown. Then the traces of cond(H) over the profile for both initial and optimized arrays are plotted. This way, the performance of the two optimization routines can be compared and the resulting arrays evaluated in turn with respect to each flight profile.

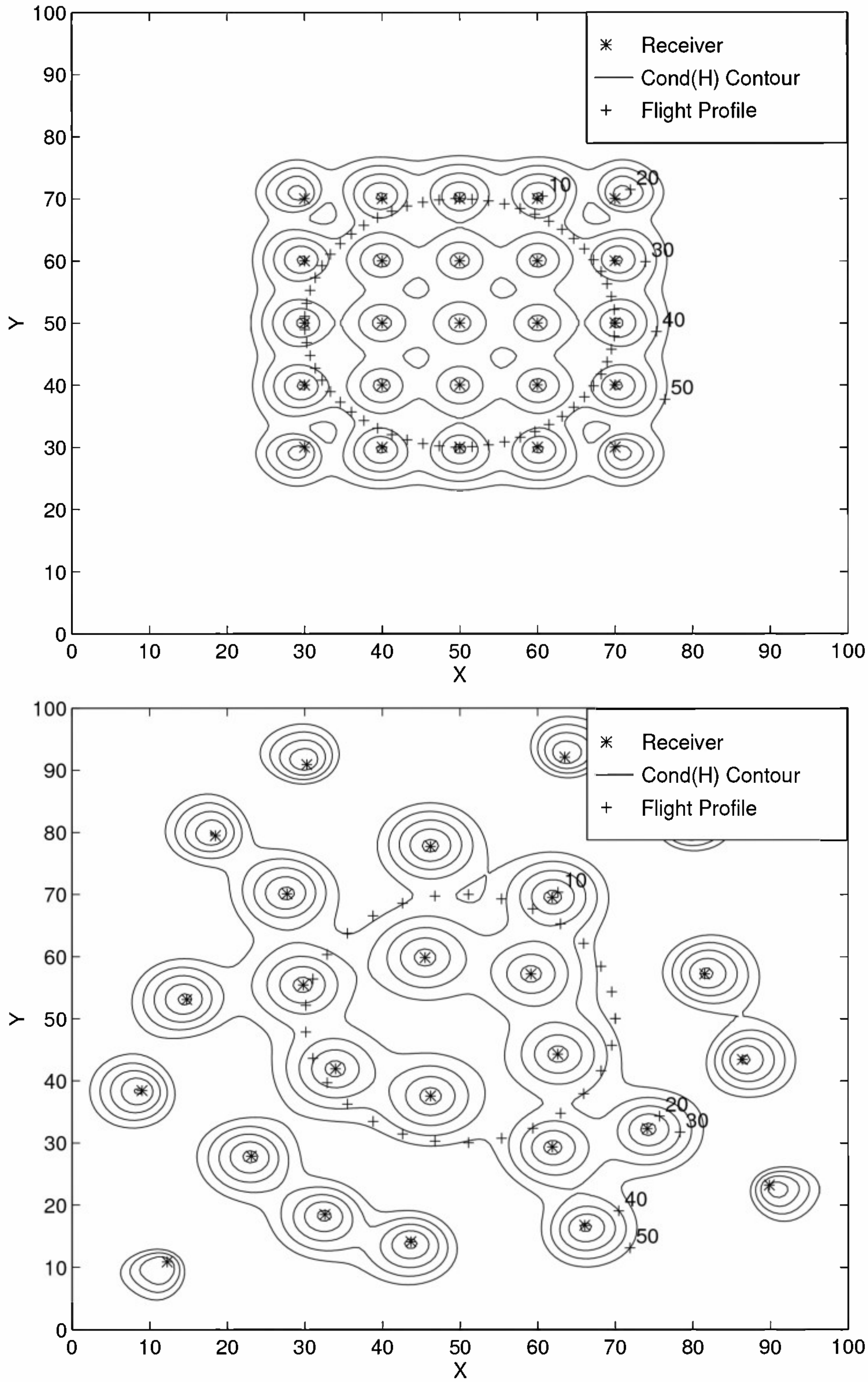


Figure 4.16 Initial Arrays, Circle Profile

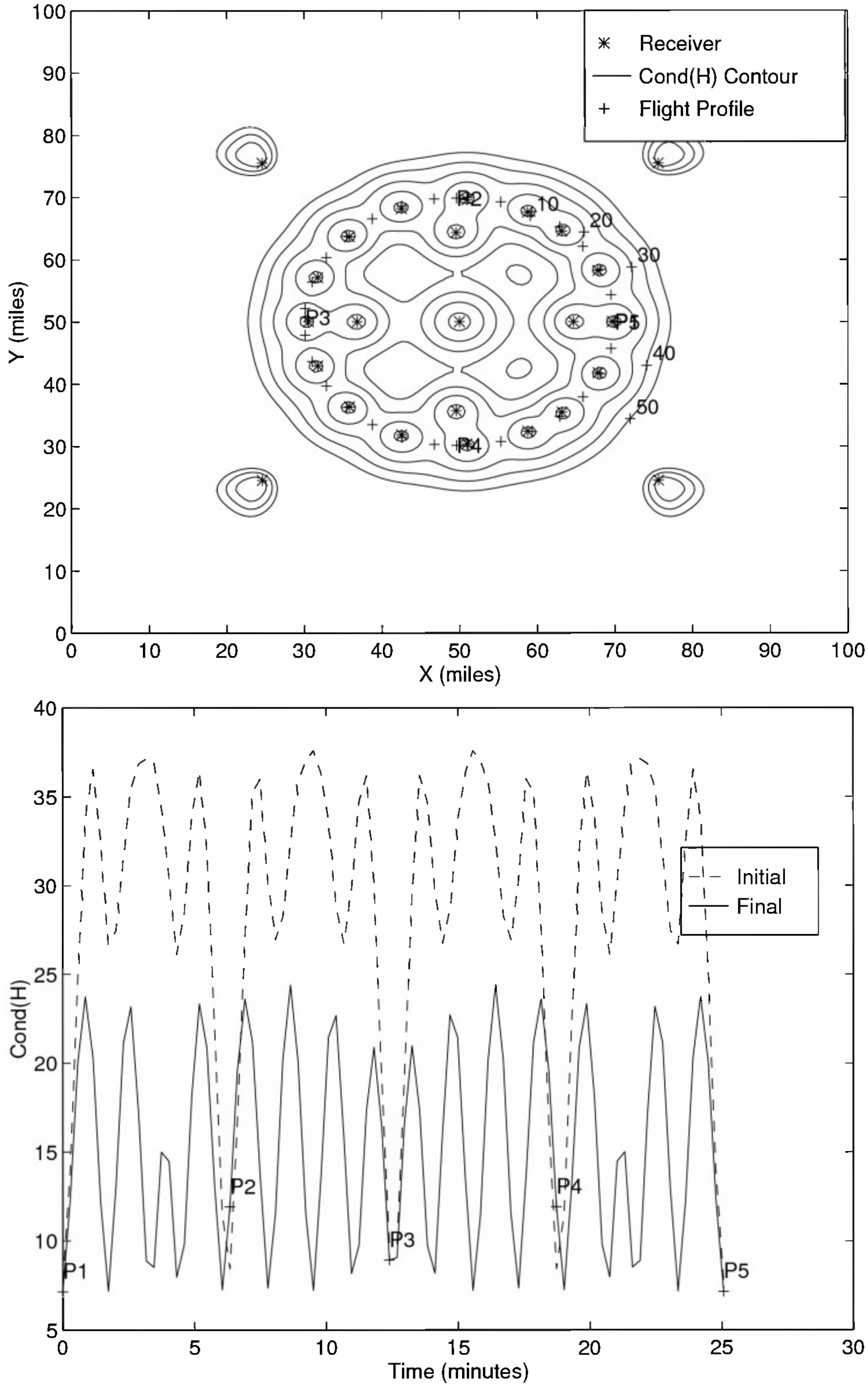


Figure 4.17 Results from constr.m with Grid array and Circle profile

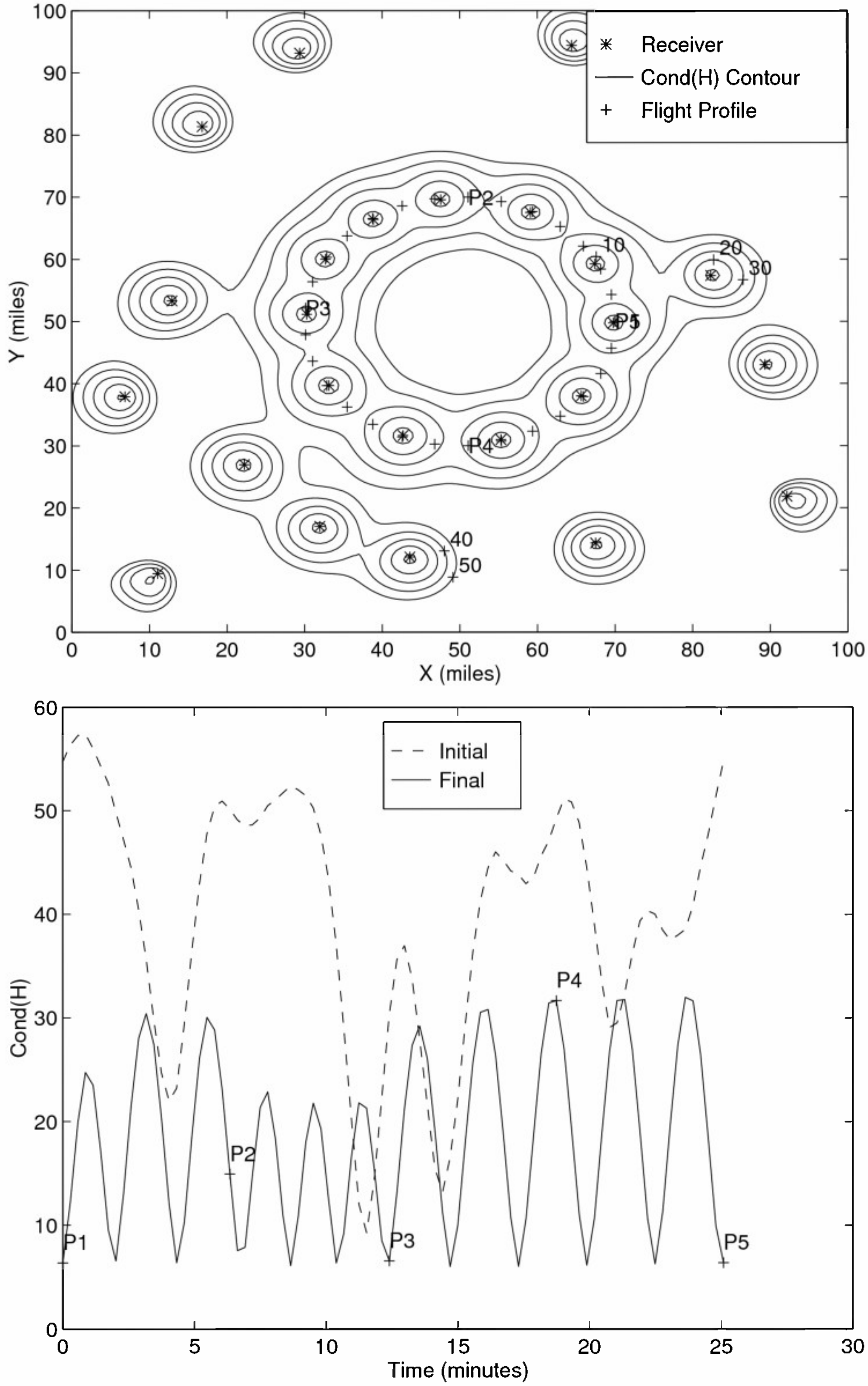


Figure 4.18 Results from constr.m with Irregular array and Circle profile

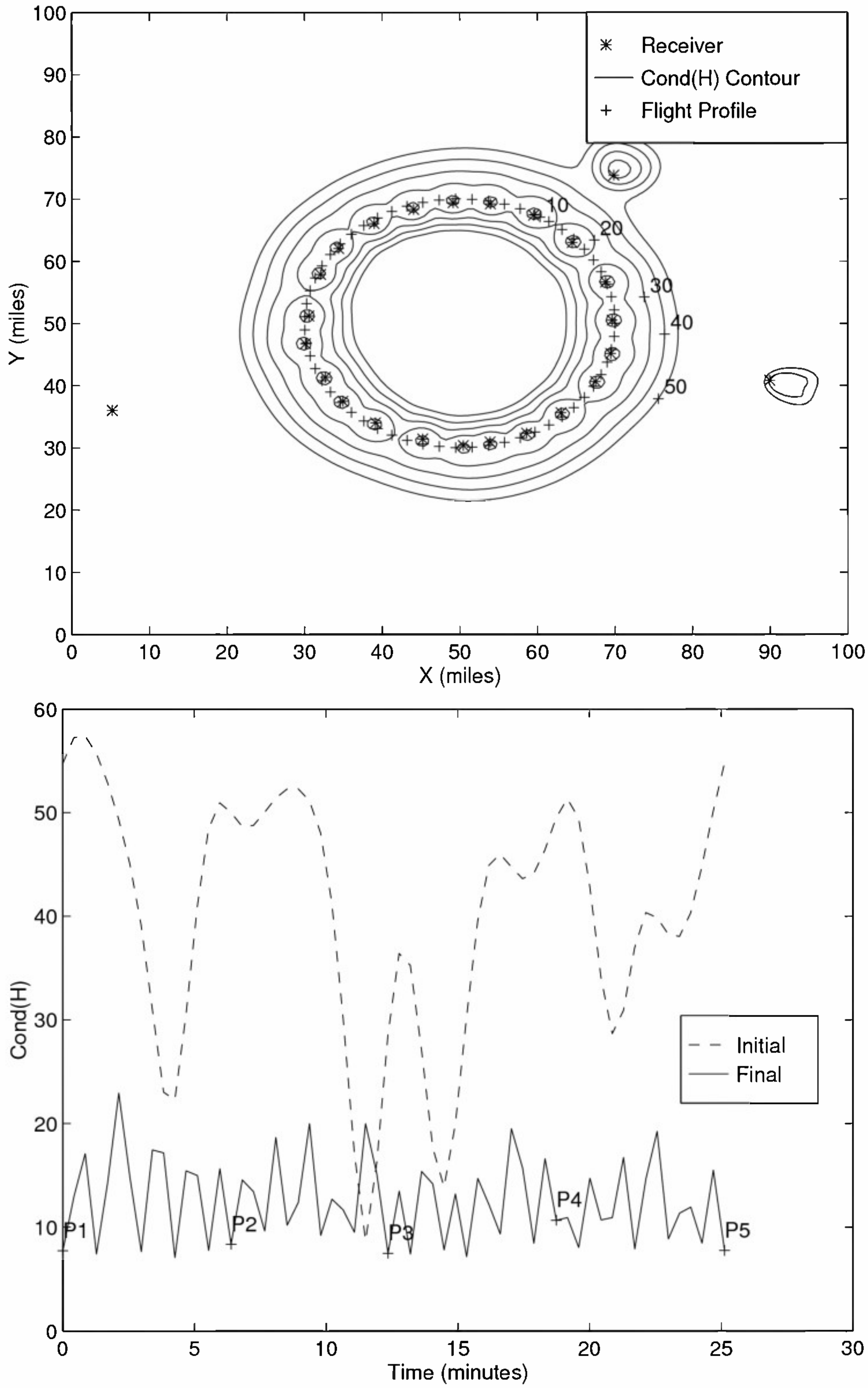


Figure 4.19 Results from M.C. Search with Circle profile

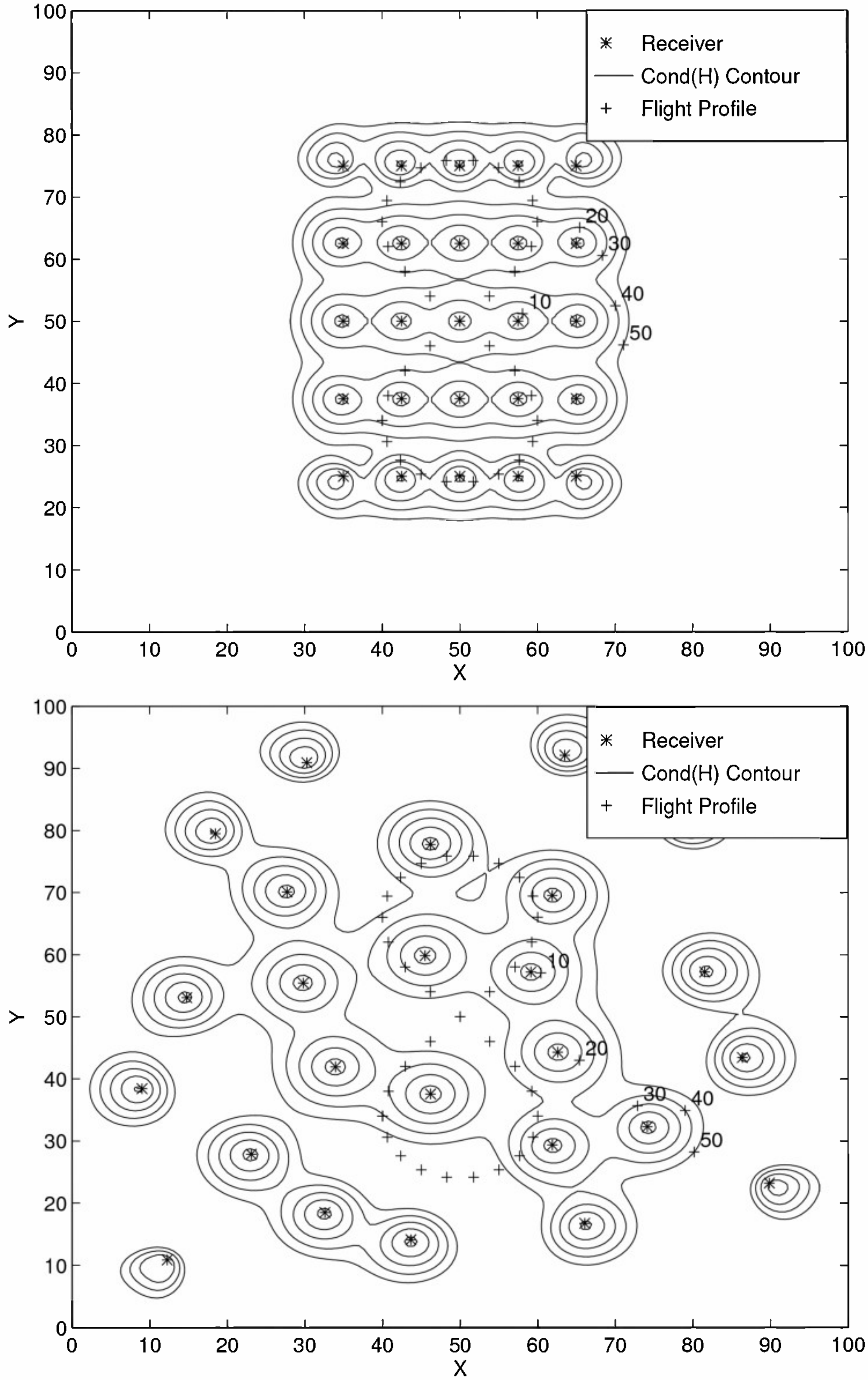
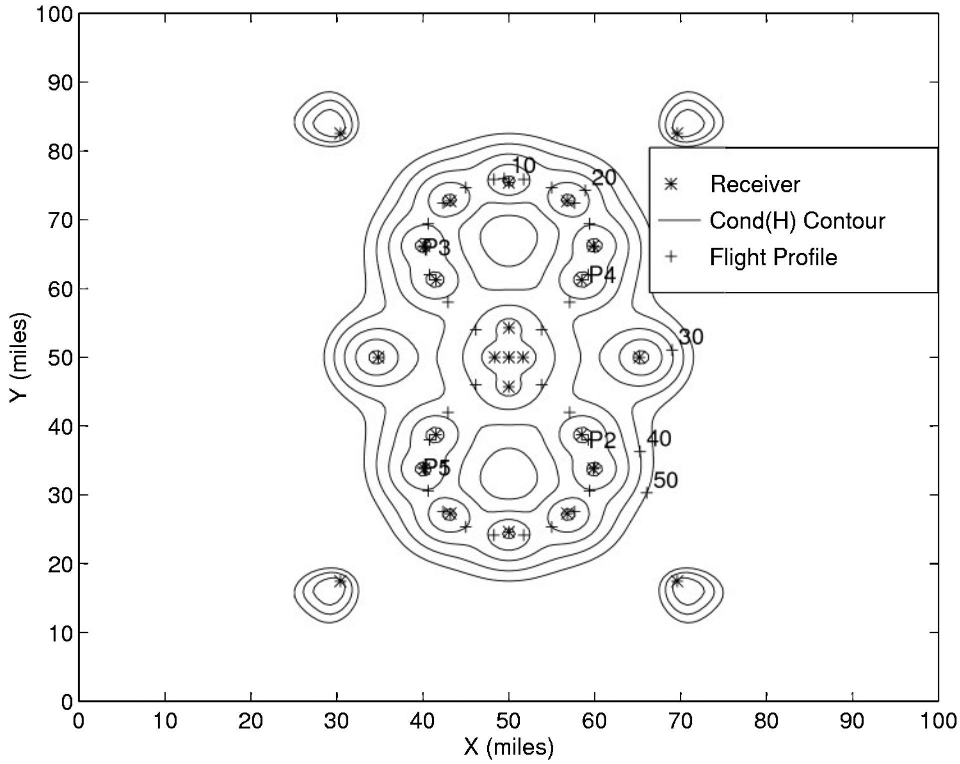


Figure 4.20 Initial Arrays, Figure Eight Profile



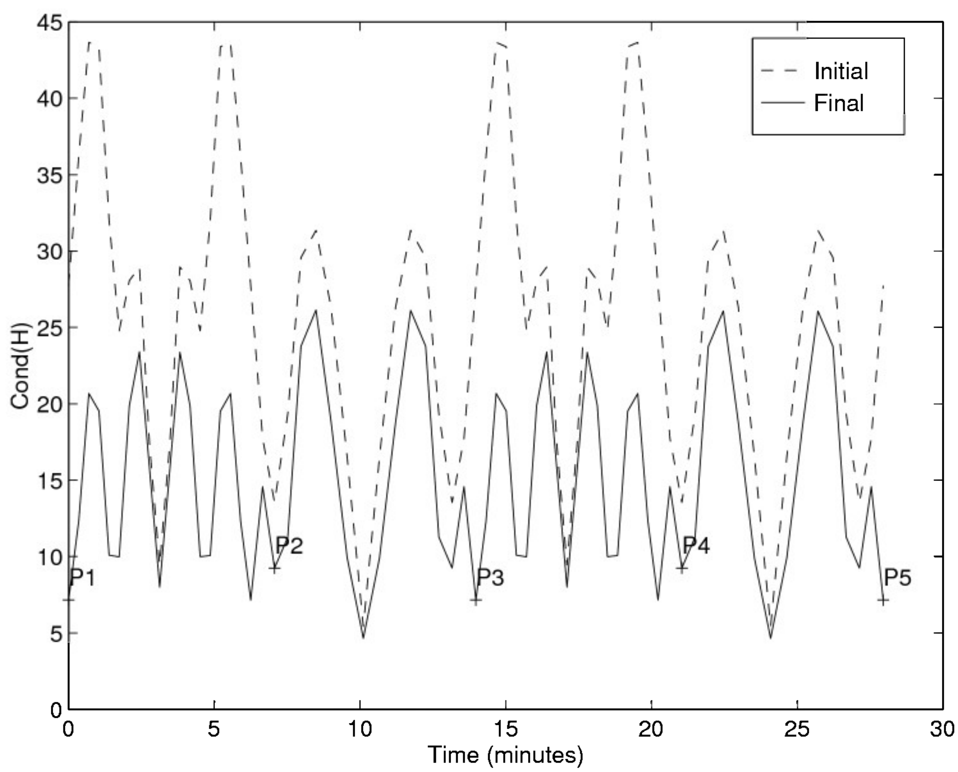


Figure 4.21 Results from constr.m with Grid array and Figure Eight profile

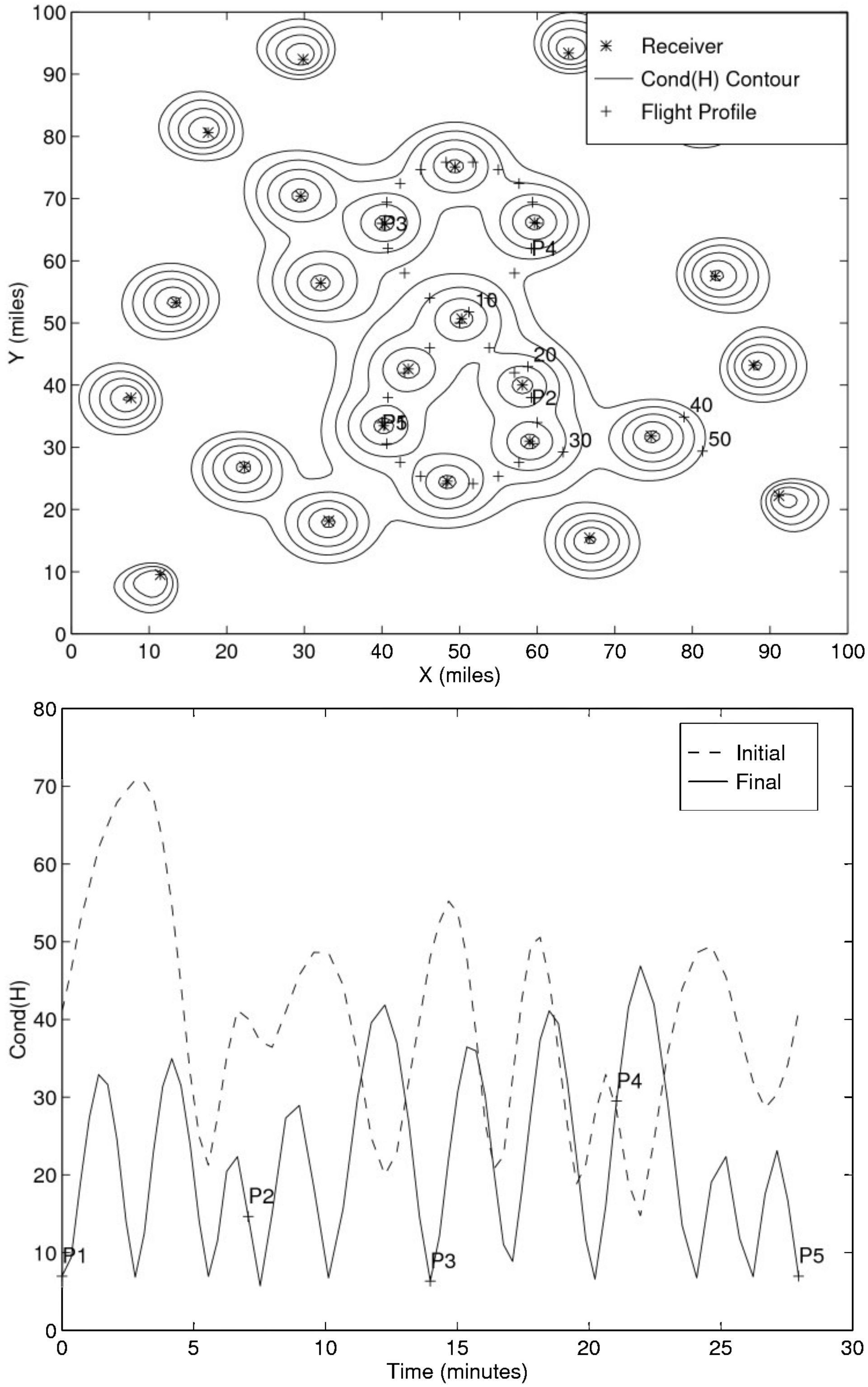


Figure 4.22 Results from constr.m with Irregular array and Figure Eight profile

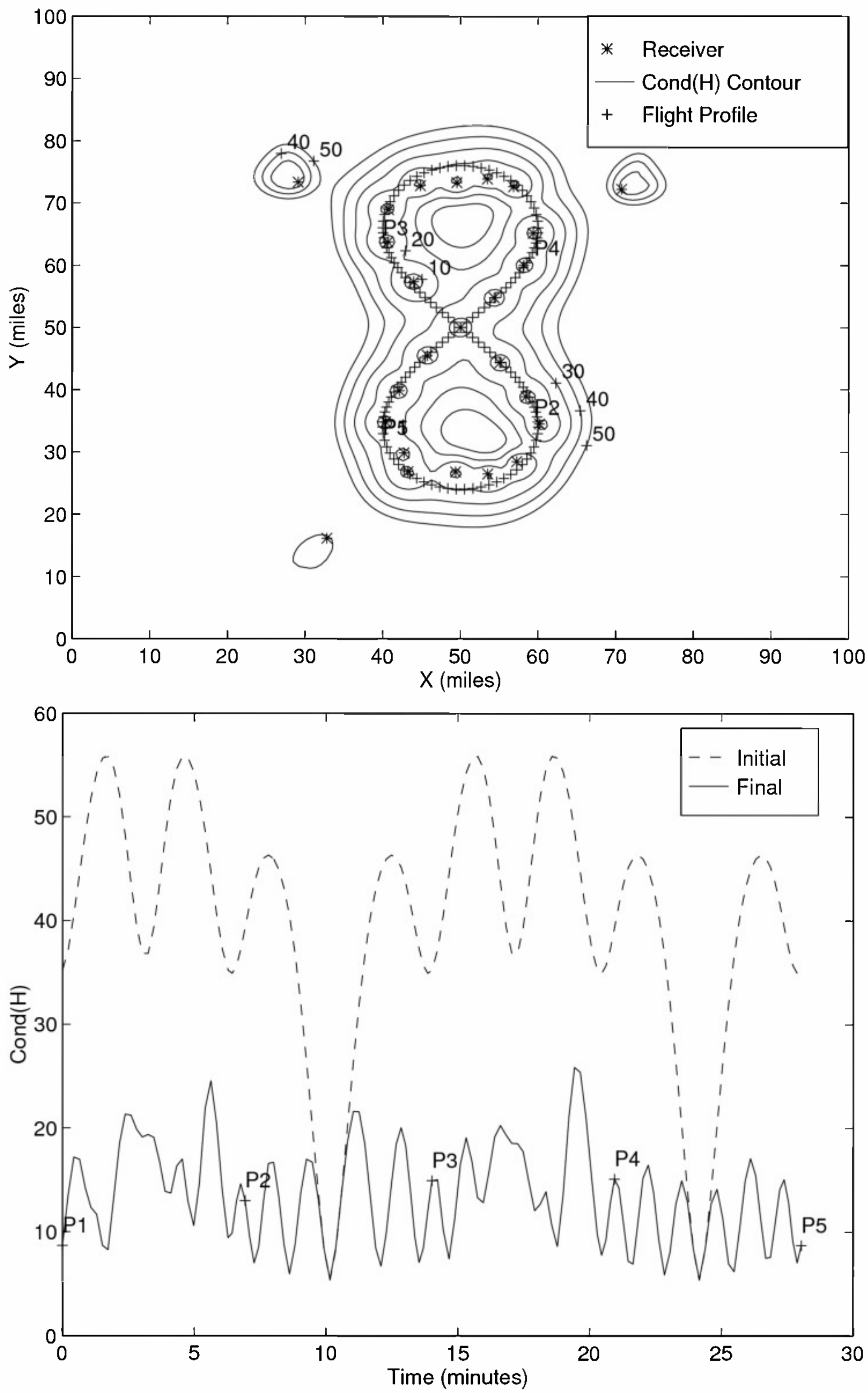


Figure 4.23 Results from M.C. Search with Figure Eight profile

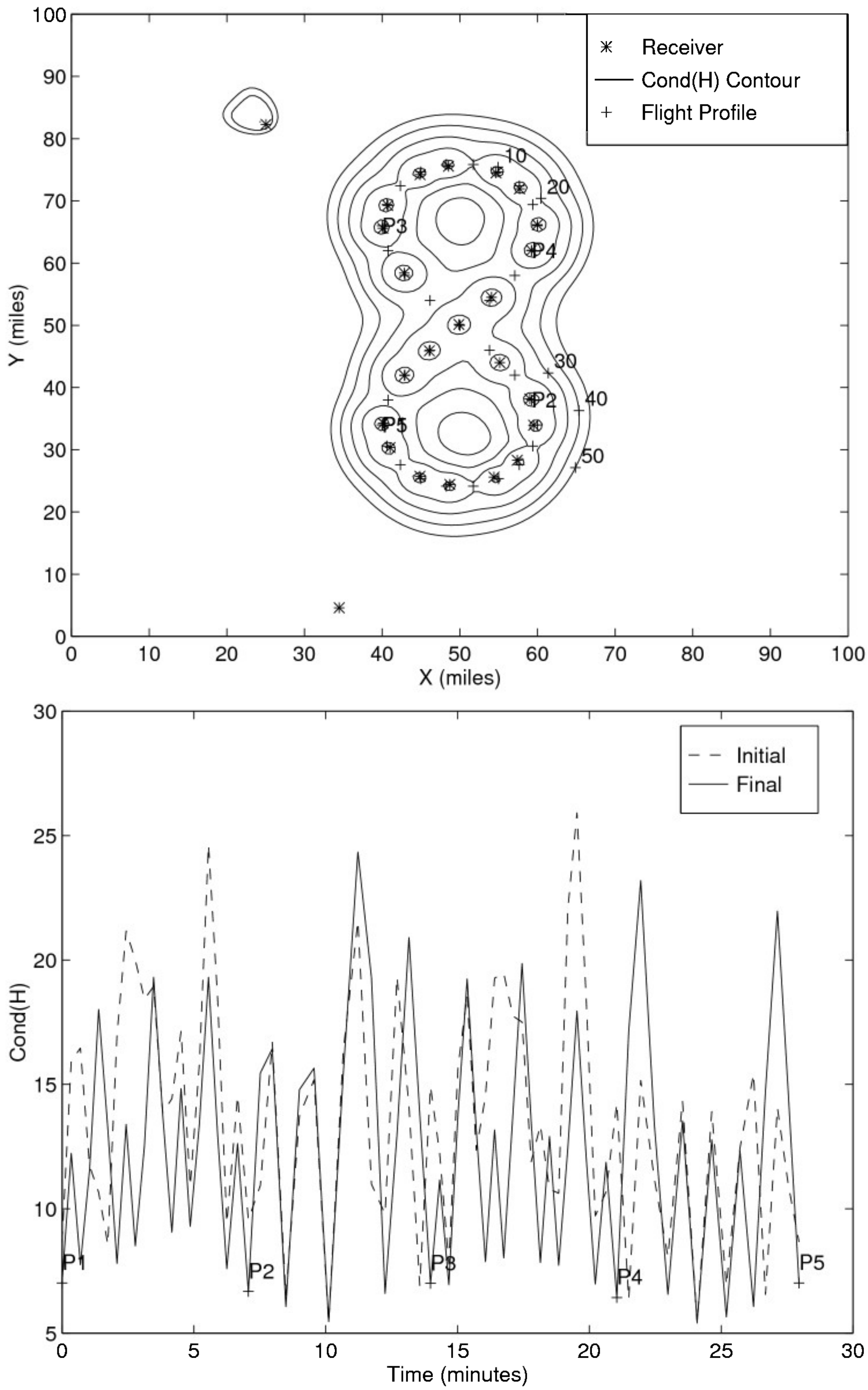


Figure 4.24 Results from constr.m with array of Fig. 4.23 and Figure Eight profile

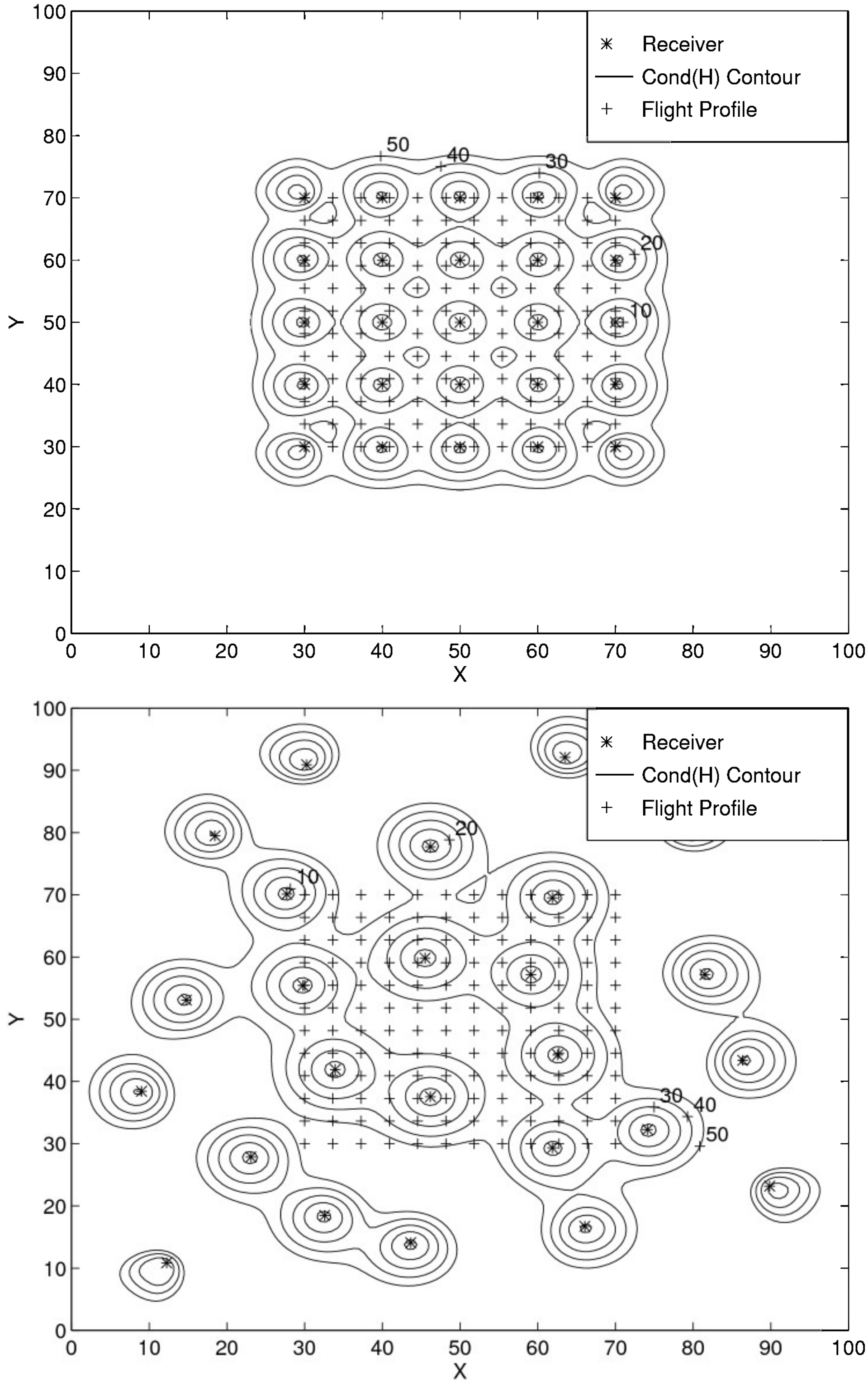


Figure 4.25 Initial Arrays, Grid Profile

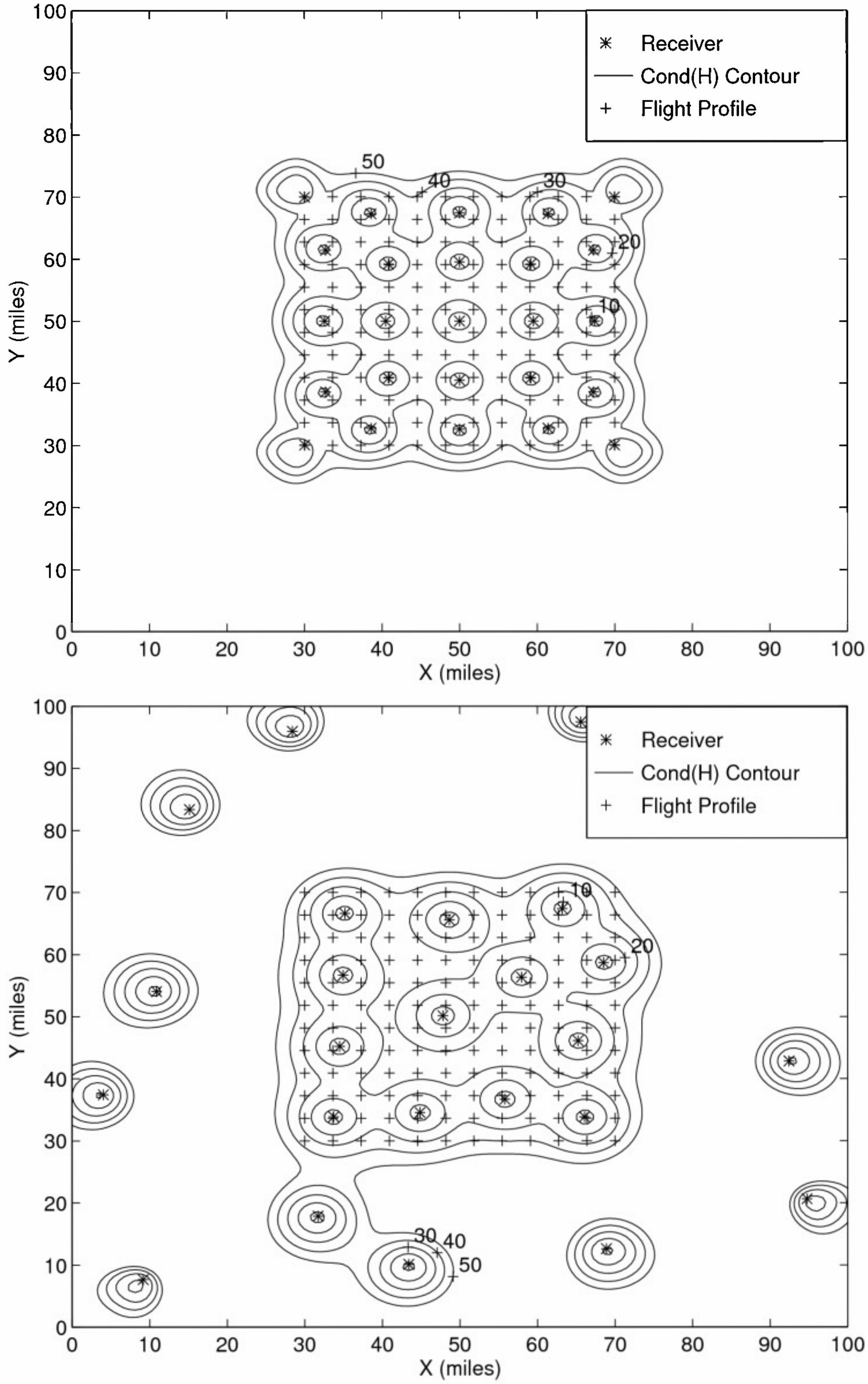


Figure 4.26 Results from constr.m with Grid, Irregular arrays and Grid profile

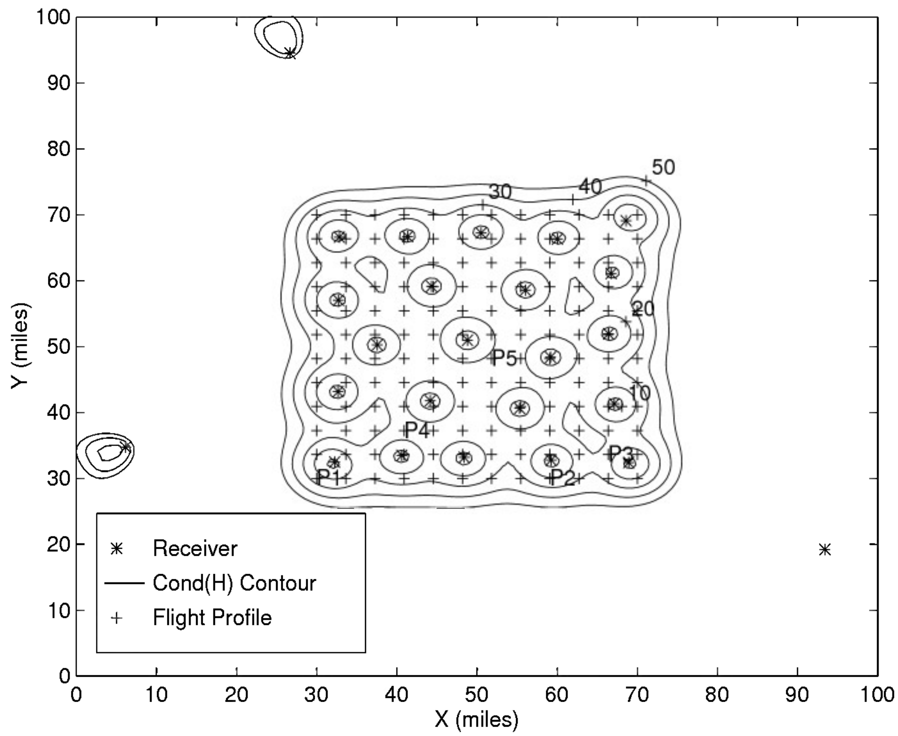


Figure 4.27   Results from M.C. Search with Grid profile

The results are summarized as follows:

1. The constr.m program is very sensitive to initial array configuration. Looking at Figs. 4.21 and 4.22, one can see that initial receiver position on the ground with respect to the flight profile makes a big difference. If the receivers are too far away from the profile, they will not be moved much in the optimization, and will likely be wasted. This means that initial arrays need to be specially chosen in order for this routine to work.

2. The Monte Carlo search program is much better at finding a good receiver array than the constr.m program, but it cannot produce as good a final array as the constr.m program can in the right conditions. As Figs. 4.16-4.19 show, the Monte Carlo search program finds a better receiver array for the circle profile than the constr.m program does, for the two initial arrays used. Given some arbitrary initial receiver array and flight profile, it is very likely that the

Monte Carlo search program would always find a better receiver array. This does not mean that the constr.m program is never worth using, however. The Monte Carlo search is very good at quickly finding a good receiver array, but the array found may not be perfect. It may have small irregularities and may have receivers in poor locations due to the search order through the array. A better array might still exist, but it likely would take moving more than one receiver simultaneously to find it, so it would never be found by using this program. However, the constr.m program does have the capability to move receivers simultaneously, and has proven to be effective in improving an array optimized to the tolerance of the Monte Carlo program. This is shown by Figs. 4.23 and 4.24. The resulting array in Fig. 4.24 is slightly better than the one in Fig. 4.23; the mean(cond(H)) over the flight profile for the receiver array of Fig. 4.24 is near eleven, and that of 4.23 is near thirteen. The array in Fig. 4.24 has regained a measure of the symmetry with respect to the flight profile which was lost in the Monte Carlo optimization, and provides a mean(cond(H)) slightly lower than that of the array in Fig. 4.23.

3. As was found in the single point planar array optimization problem, the best receiver arrays are those which have receivers directly underneath the transmitter and on the horizon at all times. Any receivers in-between those extremes do not contribute much to the array. This is best shown by the arrays optimized with respect to the circle and figure eight profiles. Note that there are a few receivers out away from the array, and the rest are located all along the flight path. This result really simplifies the design of planar arrays with respect to given flight profiles. It does show that the flight profile determines how good an array will be found, however. Therefore, the choice of flight profile is important.

4. Symmetry is a characteristic of good receiver arrays. Flight profiles should be chosen to take advantage of this. A symmetric flight profile produces a more or less symmetric array, so symmetry should be a characteristic of the flight

profile (at least the design profile), if possible. This is where the constr.m program shines. It can take a good array that is asymmetric and irregular, and fine tune it to regain symmetry. This can lower the mean cond(H) over the flight profile a couple of points, e.g., from 15 to 13, so it can be worth the time.

5. The optimized arrays are good for the flight profiles they were designed for, but little else, unless care is taken when choosing the flight profile. A circular flight profile produces a circular ring of low cond(H) at one mile. A figure eight profile produces a figure eight shaped region of low cond(H) at one mile. Of course, one could get a better low cond(H) region out of any of the optimized arrays simply by flying higher, but this may not always be desired. This is where the grid profile is useful. Although the array produced by using the grid profile in the design does not allow the cond(H) to get as low as when the circle or figure eight profiles are used, any flight that spends most of its time above the grid will be in regions of satisfactory cond(H). This versatility make the arrays designed using the grid profile (or something similar) worth considering.

The results of the array optimization effort are encouraging. It is possible to obtain a (low) mean cond(H) of less than twenty at altitudes of one mile or greater over a reasonably long flight profile (30-60 min @ 300mph) with 25 or less receivers. Although no globally optimum arrays were found, the optimization programs yielded satisfactory receiver arrays in short amounts of time. Using first the Monte Carlo program and then the constr.m program allows very good receiver arrays to be found in a matter of hours. This short optimization time allows for the possibility of re-configuring an array of mobile receivers to different flight profiles in reasonable amounts of time. If one allows the receivers to be reconfigured for every change in test profile, many different flight profiles can be handled by the system. One simply puts the new profile into the array optimizers, and out comes a satisfactory receiver array, if the flight profile is reasonable. However, if the receivers cannot be made mobile, some hard decisions need to be made. A fixed receiver array can be made to be very good for one flight profile, or can be made to be satisfactory (but not great) for many different profiles. The choice depends on which is more important, absolute accuracy or versatility. All in all, it is better to allow the receivers to be moved around in between flights, to allow for the possible reconfiguration of the receivers to handle different flight profiles with good accuracy. If this is not possible, then one must decide ahead of time what flight profiles desired and choose the receiver array configuration accordingly. If there is no reason why the minimum altitude cannot be raised to two or more miles, it should be done. Much money and time can be saved that way, without any loss of system accuracy.

V. Conclusions and Recommendations

This summarizes the results of the research to improve the accuracy of the SARS through receiver placement optimization, and discusses the key points of planar receiver array design. An array design methodology which could significantly improve the SARS' geometry is developed. More importantly, good insights into the array optimization problem are provided which yield a qualitative handle on the problem.

5.1 Gauging the Geometry Effects in GPS Positioning: Cond(H) vs. GDOP

An important result of this research is that the condition number of the H matrix from the pseudorange equation is a better measure of geometric sensitivity than the currently used GDOP metric. Although both GDOP and Cond(H) are roughly linearly correlated for a fixed number of receivers, the cond(H) is a better measure of geometric sensitivity than the the GDOP for three reasons. The first reason is because the cond(H) is independent of the number of receivers but the GDOP is not. Using the cond(H) as a measure of array configuration 'goodness' allows the geometry alone to be evaluated, regardless of the number of receivers in the array. The second reason is that the cond(H) provides a maximum bound to the amplification of the error possible when solving the pseudorange equations for position. This provides a hard upper limit on error amplification, which is a useful quantity to have when designing a GPS like system's geometry to maximize accuracy. The third reason is that cond(H) allows the sequential programming optimization routine to produce better arrays than if the GDOP were used. Hence, we used cond(H) as a measure of the geometric effect of pseudorange measurement error on the GPS position fix in general, and with respect to the SARS application in particular. It is interesting to note that this modeling step also helps the optimization algorithm. The algorithm does not have as much difficulty with convergence to undesired local minima in the cost function when the optimization criterion is based on cond(H), so it has a better chance of producing a useful receiver array. These reasons show that the condition number of the H matrix makes a more useful and precise measure of the impact of GPS measurement geometry on the navigation solution's accuracy than the commonly used GDOP metric.

5.2 Array Optimization

The accuracy of the SARS, or any ground based inverted GPS system, can be significantly improved by numerically optimizing the design of the receiver array with respect to the desired flight profile of the transmitter, using the optimization tools developed in this research: the Monte Carlo search program adapted to this problem, and the conventional Sequential Quadratic Programming optimization using constr.m from the MATLAB Optimization Toolbox. This optimization yields receiver arrays which cause regions of airspace to have reduced "GDOP." This lowers the sensitivity of the pseudorange equations to measurement errors, hence improving the accuracy of the computed position. Although the arrays designed by the optimization programs are not guaranteed to be globally optimum, they are quite good, providing satisfactory receiver arrays for use in the SARS. Moreover, the optimization tools forged in this research can be applied to any proposed receiver array so that an improved array ensues.

The best results come from first running the Monte Carlo optimization program to get the receiver array close to the global optimum, then running the sequential quadratic programming routine based on constr.m to search for the local optimum. This procedure consistently finds arrays that are just about as good as the design specifications and constraints allow, *regardless* of the initial array configuration input to the Monte Carlo search program. This composite array optimization can be done in four to eight hours (of unsupervised run time) on a Sparc 10 or Sparc 20 workstation , enabling re-configuration of the receiver array for different flight profiles to be done within a reasonable amount of time. Actually, moving the receivers might take longer, depending on the number of people in the work crew.

5.3 Key Array Design Points

The results of the array optimization research show that there are several key points in the design of receiver arrays for inverted GPS systems. These are:

1. *Fly as high as possible.* The main problem with ground based arrays intended for use with aircraft is that most of the region of good geometry lies well above the maximum altitudes of most aircraft. Without adding additional measurements or airborne receivers, there is little that can be done about this problem, except fly the aircraft as high as possible.

2. *The best planar array with four receivers is three receivers on the horizon and one directly underneath the transmitter.* Adding more receivers to the array will not improve the geometry, but neither will the additional receivers detract from the geometry, as long as the array is re-optimized for the greater number of receivers. Optimum arrays with large numbers of receivers tend to be superpositions of optimum arrays with smaller numbers of receivers. This concept of superposition can be applied to the optimization problem with multiple transmitter points. If there are three or more receivers spaced far apart on the horizon and one receiver below each transmitter point on the flight profile, the receiver array will have very good geometry, at least at the transmitter points designed for. Therefore, one way of quickly coming up with a good initial array to feed the optimization routines is to put three receivers out at the border of the test range, then put the rest of the receivers, equally spaced, along the ground directly underneath or close to the desired flight profile.

3. *Tradeoff 1* There is a tradeoff between minimum transmitter altitude and receiver spacing in the array. The lower the altitudes at which the transmitter

will be flown, the closer together the receivers need to be in the portion of the array directly underneath the flight profile to produce the same cond(H). Unfortunately, the good range of aircraft altitudes for a given receiver spacing depends on the specific receiver array and flight profile used. It varies enough (even between parts of the same profile) that no simple table can be given. An easy way to work the tradeoff is to hand-pick arrays (as discussed above) for the flight profile, starting with few receivers, evaluating the cond(H) trace along the profile, and keep adding receivers to the projection of the flight profile on the ground (the ground track) until a cond(H) trace close to being satisfactory is found. If this results in too many receivers being used, the flight profile probably needs to be modified. Possible modifications to the flight profile are raising the altitude range of the flight profile, making the flight shorter, or looping the flight back upon itself (as in a circle or figure eight).

4. *Tradeoff 2* There is a tradeoff between performance and versatility for planar receiver arrays. For planar arrays, generally speaking, the better an array performs for one flight profile, the worse it performs for flights that are different from the one it was designed for (because there need to be lots of receivers underneath the flight profile). This makes no difference if only one flight profile will be used, or if the receivers can be easily moved. But it does make a difference if it is desired for several flight profiles to be used simultaneously, or if the receivers cannot be moved easily. Two ways could be used to design a receiver array for simultaneous use with several flight profiles. The first way is to optimize the array with respect to a flight profile composed of all the desired flight profile points simply added together into one long list. For this to work well, the flight profiles had better have significant overlap, or many receivers will be needed to keep them from being spread too thin. The other way is to simply design for an airspace of low cond(H) by filling a region at minimum altitude with evenly spaced transmitter points, like the grid profile

shown in chapter four. This technique will produce a contiguous, three dimensional airspace of low cond(H) that can be used for any flight profile that will fit within it, making room for aerobatic maneuvers that might not fit into arrays optimized about a single flight profile. Both techniques have their uses, depending on the design requirements and constraints.

5. *Altitude differences in receiver locations due to terrain do not appreciably affect the results.* Optimizations run using the mountainous terrain at the White Sands Missile Range (2500 ft altitude differences) showed that the best receiver locations were invariably determined by the flight profile, not the topography. The topography produces a third order effect in the cond(H). Adding several receivers atop 2500 foot towers to a completely planar array brought the cond(H) at ground level from infinity to several hundred, but only changed the tenths place on the cond(H) at one mile above the ground. For a real topography, the changes would be somewhat less striking. The higher the aircraft flies, the less important altitude variation in the receivers' locations is. In other words, at altitudes high enough to produce satisfactory cond(H) traces for the flight profile in the first place, the effects of realistic altitude variations from terrain are insignificant. It would take gross altitude differences in receiver placement (miles!), to significantly improve the cond(H) as compared to the completely planar array. These gross altitude differences cannot be caused by terrain or any kind of structure built on the ground.

6. *The masking of receivers from the transmitter line of sight does not significantly impact the geometric sensitivity.* Line of sight (LOS) difficulties that significantly affect the cond(H) have not occurred at flight altitudes of one mile or greater in this research. Although transmitter-receiver line of sight may be broken for as many as five or six receivers out of the 25 receiver array over the course of the flight profile, the cond(H) is not significantly increased. After all, the research has shown that one needs only three receivers on the horizon and

one receiver directly underneath the transmitter to have good geometry. The receivers that are directly underneath the transmitter will not lose LOS, and of the rest, only three of them are really needed to maintain good geometry.

7. *LOS can be a constraint in the optimization, but it takes so much time that it is better to use common sense to prevent loss of LOS.* If LOS needs to be maintained without breaks through as much of the flight as possible for as many receivers as possible, precautions need to be taken. The array optimization programs do not take LOS into account. It was tried several times over the course of this research, and was found to not be worth the incredible time (weeks!) it took to run the optimizations. When the geometry of the arrays optimized with the LOS check in use proved identical to those found without considering LOS, the LOS check was discarded. Therefore, LOS may be considered after the fact. Fortunately, checking the LOS over a flight profile represented by about one hundred to two hundred points takes only several minutes, so an LOS check on the finished array and flight profile can be done, to see if either array or flight profile should be modified to prevent loss of LOS. Most likely, all that will be required is small changes that do not require that the array be re-optimized. For example, moving the receivers as much as a mile or so, to take advantage of local hilltops or existing buildings/towers does not significantly change the geometry of the array, but does allow LOS to be maintained much more easily. Of course, the flight profile could be chosen to occur over a large flat area or basin where LOS would not be a problem anyway, and all the difficulty would be avoided. For the SARS, the map of the missile range shows that there is indeed a large basin surrounded by mountains inside the range. Making the flight profile fly over the basin, instead of the mountains, would go a long way towards eliminating LOS concerns.

5.4 A Technique for Improving the SARS' Geometry

The research indicates that the main difficulty with a ground based array is that it takes too many receivers to ensure that one is always directly underneath the transmitter. If this need for a receiver directly underneath the transmitter could be eliminated, many fewer receivers would be needed in the array to produce the same values of cond(H). Also, the dependence of the array configuration on the flight profile would be significantly reduced, allowing more freedom in choosing flight profiles, and ending the necessity for array re-configuration for a different flight profile. This improvement can be achieved by allowing a receiver to be airborne, attached to a second aircraft flying directly above the test aircraft, as high as possible. A balloon could be used instead, but since it could not keep up with the test aircraft, it would need to be at an *extremely* high altitude. Of course, there appears to be one problem with this proposed scheme: the mobile receiver needs to have its position known very accurately in order for this technique to work. This problem is easily overcome by equipping both aircraft with receivers, thus achieving /em self calibration. In this case, one runs two flight tests simultaneously at the price of one flight test, the proverbial "two for one." This concept is illustrated in Fig. 5.1.

Hence, the solution to this problem is to allow for the position of the mobile receiver to be found in the solution process. This is accomplished by placing a transmitter on the same craft as the mobile receiver (and shielding the mobile receiver from it). This configuration provides pseudorange measurements from both aircraft to the ground based receivers, and one pseudorange measurement between the transmitter on the first aircraft and the mobile receiver on the second aircraft (or balloon). The distance between the second transmitter and the mobile receiver (both on the same craft) is measured on the ground. Now, there are enough measurements to solve for both the position of both transmitters and the mobile receiver. This will dramatically improve the measurement geometry of the system, allowing

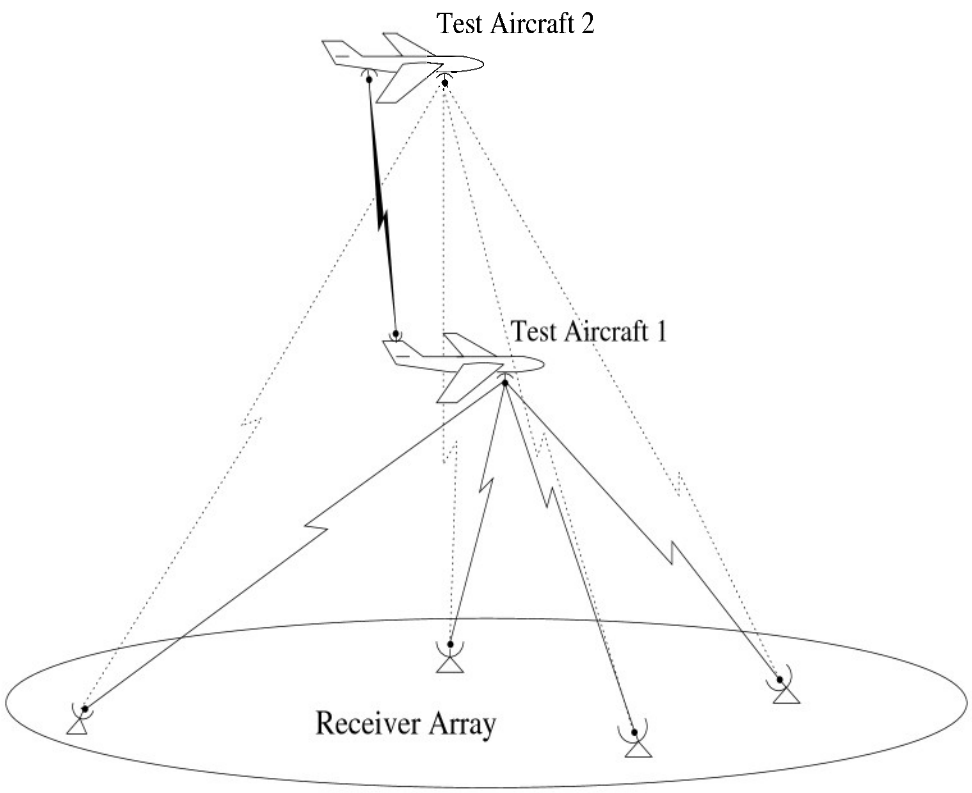


Figure 5.1   A Technique for Improving the SARS' Geometry

higher accuracy with fewer receivers (e.g., five or six instead of twenty five) than in a purely ground based array. Of course, this method requires two aircraft, but there is no reason that two tests could not be performed simultaneously, so it need not be wasteful.

This dramatic improvement is shown by the following example. The number of receivers in the array are just four: one at the center and three on the border of the test range. The figure eight profile from Fig. 4.14 is used. One aircraft flies at an altitude of one mile above ground level; the other flies directly above the first aircraft at an altitude of four miles above ground level.

Fig. 5.3 shows that a maximum cond(H) of 5.6 is obtained over the course of the flight. This is an incredibly small number compared to the cond(H) over the same flight profile for the plain SARS without the second aircraft and receiver. The amount of equipment needed to attain this level of performance is two transmitters, two aircraft, and six receivers.

Although the best configuration is when one aircraft is directly above the other, there is significant leeway in the relative positioning of the aircraft. The line of sight vector between the two aircraft should be as close to perfectly vertical as possible, to attain the best geometry, but the alignment of the aircraft can withstand significant variations and still provide good geometry. Figures 5.4 and 5.5 show a case where the aircraft at an altitude of four miles is always five miles north of the other aircraft. Although the geometry is not as good as the case in Fig. 5.3, it still is very good, much better than is possible without using both aircraft. This shows that significant variation in relative aircraft positioning can be handled by this self calibration technique. Therefore, no precision flying is needed.

In addition, this technique allows much leeway in choosing a good flight profile. As long as the flight stays mostly within the triangle formed by the three receivers on the boundary, the geometry will be good. In theory, the pilots could fly the aircraft in any flight profiles they choose, as long as the aircraft stay over the test range, keep their altitude separation, and don't get more than five or so miles apart horizontally. This technique, although it requires an extra transmitter and an additional aircraft, solves the low altitude geometry problem for the SARS, allowing accuracy as good or better than from the GPS satellites to be obtained from the SARS, without strict constraints on the flight profile. If very accurate low altitude tests are desired, this method should be considered.

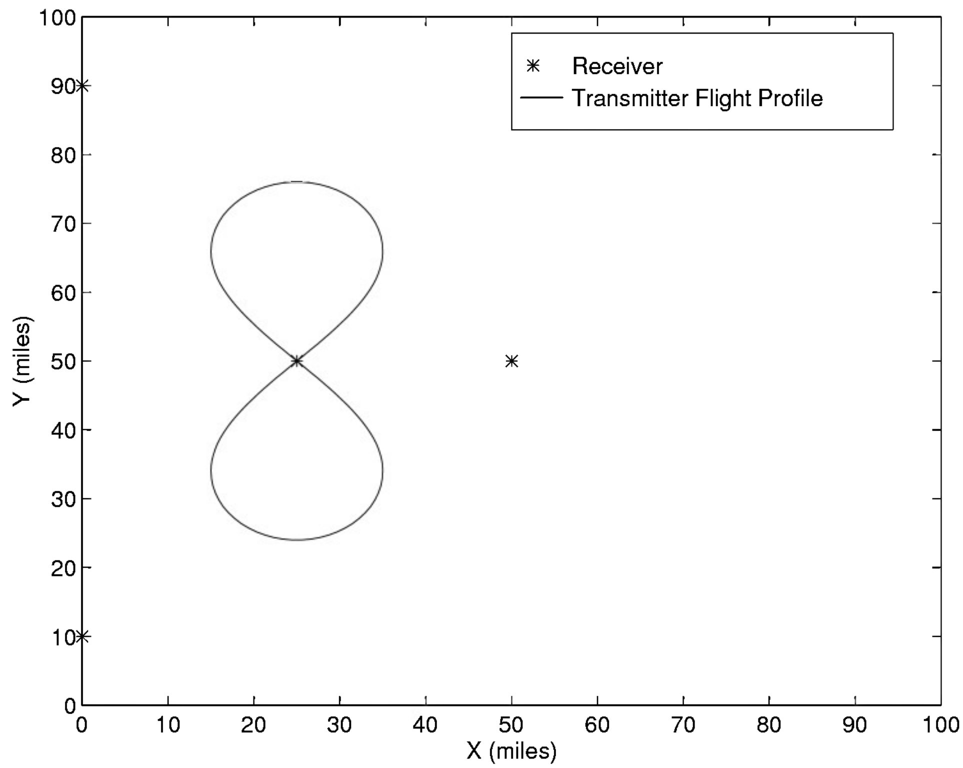


Figure 5.2 SARS Array and Flight Profiles for the Self Calibration Technique

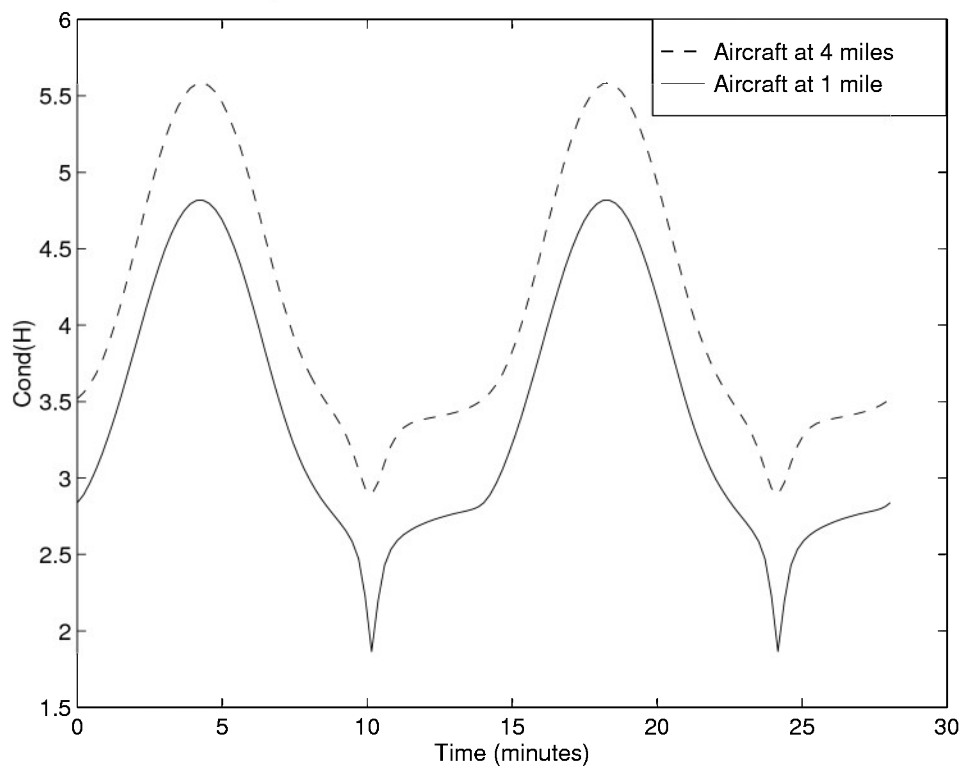


Figure 5.3 Cond(H) for Both Aircraft Over the Flight Profiles

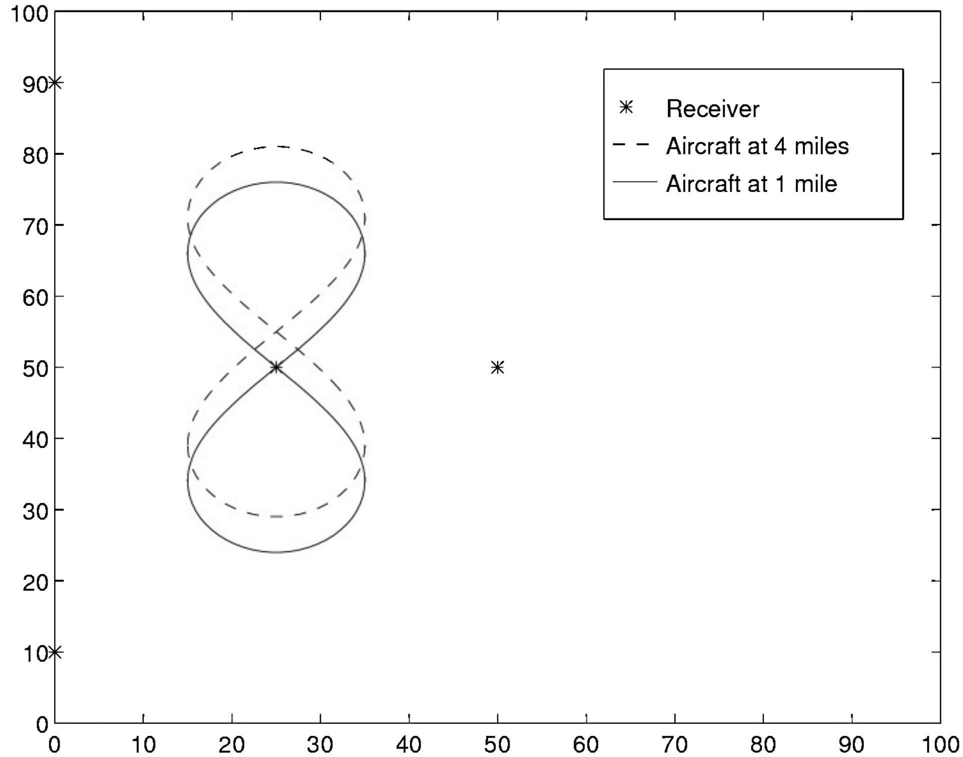


Figure 5.4 SARS Array and Flight Profiles for the Self Calibration Technique

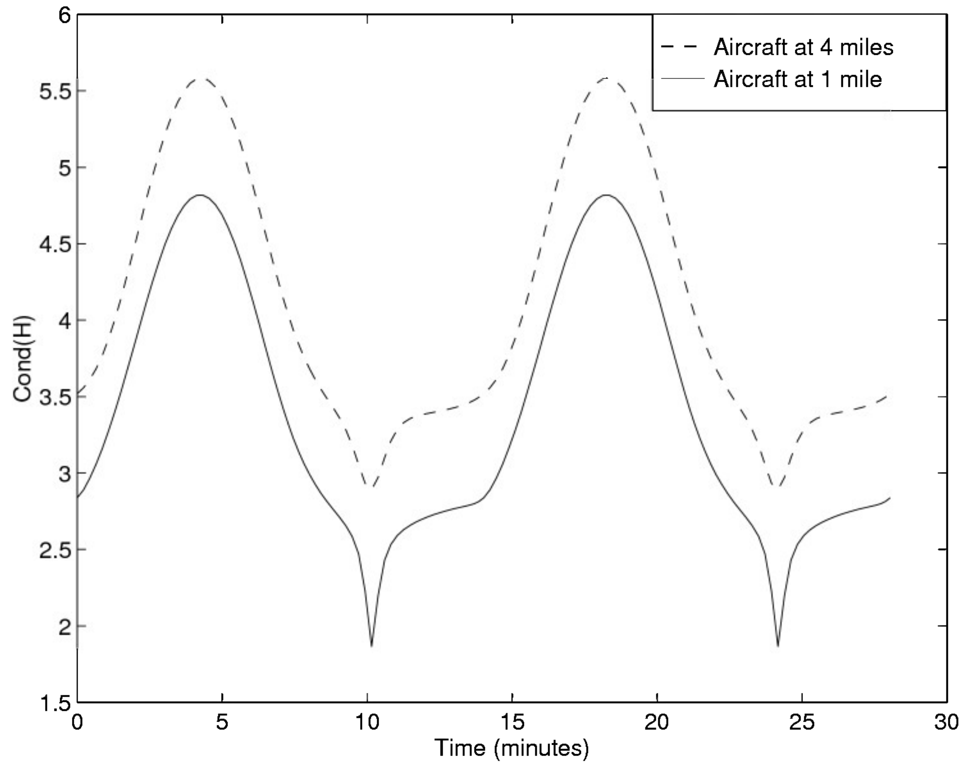


Figure 5.5 Cond(H) for Both Aircraft Over the Flight Profiles

Bibliography

1. Bancroft, S. "An Algebraic Solution of the GPS equations," *IEEE Transactions on Aerospace and Electronic Systems*, *21*(1):56–59 (November 1985).
2. Bohenek, Brian J. *The Enhanced Performance of an Integrated Navigation System in a Highly Dynamic Environment*. MS Thesis, AFIT/GE/ENG/94D-01, School of Engineering, Air Force Institute of Technology, Wright-Patterson AFB, OH, December 1994.
3. Canadian GPS Associates. *Guide to GPS Positioning*. Fredericton, N.B., Canada: University of New Brunswick Graphic Services, 1987.
4. Chaffee, J. and J. Abel. "On the Exact Solutions of Pseudorange Equations," *IEEE Transactions on Aerospace and Electronic Systems*, *30*(4):1021–1030 (October 1994).
5. Cox, Jr, D. "Integration of GPS with Inertial Navigation Systems," *Navigation, Journal of the Institute of Navigation*, 144–152 (1978).
6. Hansen, Neil P. *Incorporation of Carrier Phase Global Positioning System Measurements into the Navigation Reference System for Improved Performance*. MS Thesis, AFIT/GE/ENG/93D-40, School of Engineering, Air Force Institute of Technology, Wright-Patterson AFB, OH, December 1993.
7. Hebert, Jeffrey M. *Velocity Determination for an Inverted Pseudolite Navigation Reference System*. MS Thesis, AFIT/GE/ENG/95D-06, School of Engineering, Air Force Institute of Technology, Wright-Patterson AFB, OH, December 1995.
8. Hwang, Patrick Y.C. "Kinematic GPS for Differential Positioning: Resolving Integer Ambiguities on the Fly," *Navigation, Journal of the Institute of Navigation*, *4*:205–219 (1993).
9. Lachapelle, G. et al. "High Precision GPS Navigation with Emphasis on Carrier Phase Ambiguity Resolution," *Marine Geodesy*, *15*(4):253–269 (1992).
10. The MathWorks, Inc., Natick, MA. *MATLAB (registered trademark)* (Version 4.2c Edition), November 1994.
11. Maybeck, Peter S. *Stochastic Models, Estimation, and Control*, *I*. New York: Academic Press, Inc., 1979. Republished, Arlington, VA: Navtech, 1994.
12. Maybeck, Peter S. *Stochastic Models, Estimation, and Control*, *II*. New York: Academic Press, Inc., 1982. Republished, Arlington, VA: Navtech, 1994.
13. Maybeck, Peter S. *Stochastic Models, Estimation, and Control*, *III*. New York: Academic Press, Inc., 1982.
14. Milliken, R. J. and C. J. Zoller. "Principles of Operation of NAVSTAR and System Characteristics," *Navigation, Journal of the Institute of Navigation*, *25*(2):3–14 (Summer 1978).

15. Mosle, William B. *Detection, Isolation, and Recovery of Failures in an Integrated Navigation System*. MS Thesis, AFIT/GE/ENG/93D-28, School of Engineering, Air Force Institute of Technology, Wright-Patterson AFB, OH, December 1993.

16. Negast, William J. *Incorporation of Differential Global Positioning System Measurements Using an Extended Kalman Filter for Improved Reference System Performance*. MS Thesis, AFIT/GE/ENG/91D-41, School of Engineering, Air Force Institute of Technology, Wright-Patterson AFB, OH, December 1991.

17. Raquet, J. et al. "Development and Testing of a Mobile Pseudolite Concept for Precise Positioning," *Proceedings of ION GPS-95*, 1–11 (September 1995). Palm Springs.

18. Robbins, James E. "Reference Trajectories from GPS Measurements," *Navigation, Journal of the Institute of Navigation*, *35*(1):89–103 (Spring 1988).

19. Schwarz, K. et al. "A Comparison of GPS Kinematic Models for the Determination of Position and Velocity along a Trajectory," *Manuscripta Geodaetica*, *14*:345–353 (1989).

20. Siouris, George M. *Aerospace Avionics Systems: a modern synthesis*. San Diego: Academic Press, Inc., 1993.

21. Stacey, Richard D. *A Navigation Reference System (NRS) Using Global Positioning System (GPS) and Transponder Aiding*. MS Thesis, AFIT/GE/ENG/91M-04, School of Engineering, Air Force Institute of Technology, Wright-Patterson AFB, OH, March 1991.

22. Tang, Wang, et al. "A Kinematic Carrier Phase Tracking System for High Precision Trajectory Determination," *Proceedings of ION GPS-94* (September 1994).

23. Vasquez, Juan R. *Detection of Spoofing, Jamming, or Failure of a Global Positioning System (GPS)*. MS Thesis, AFIT/GE/ENG/92D-37, School of Engineering, Air Force Institute of Technology, Wright-Patterson AFB, OH, December 1992.

Vita

PII Redacted

Jason Bryce McKay was born on [redacted] in [redacted]. Upon graduation from Oakton High School (Fairfax, VA) in 1990, Jason began his undergraduate studies at the University of Illinois, Urbana. While at the University of Illinois, Jason enrolled in the Air Force Reserve Officer Training Corps (ROTC). Four months before graduation, Jason met his future wife, Miss Kristen Lock. After the 'five year plan' was completed, Jason graduated with honors from the U of I in May of 1995. After graduation, Jason was commissioned into the United States Air Force. On 5 June 1995, Jason reported to the Air Force Institute of Technology (AFIT) for his first assignment, graduate school, leaving Kristen behind in Illinois. This would not do, so Jason and Kristen were married next year in June, between AFIT quarters. While at AFIT, Jason worked on a Masters of Science Degree in Electrical Engineering with a concentration in navigation, guidance, and control. Upon graduation from AFIT in December 1996, Jason reports to the Laser Applications Branch of the Electro-Optics Division, Wright Labs.

PII Redacted

Permanent address: [redacted]